COMPARING THE PERFORMANCE OF WIND SPEED DISTRIBUTIONS: THE CASE OF CENTRAL AND SOUTHERN IRAQ

OTHMAN SAFAA AL-SHALASH

Master's Thesis

Graduate School
Izmir University of Economics
Izmir
2020

COMPARING THE PERFORMANCE OF WIND SPEED DISTRIBUTIONS: THE CASE OF CENTRAL AND SOUTHERN IRAQ

OTHMAN SAFAA AL-SHALASH

A Thesis Submitted to

The Graduate School of Izmir University of Economics
Master's Program in Applied Statistics

Izmir

2020

ABSTRACT

COMPARING THE PERFORMANCE OF WIND SPEED DISTRIBUTIONS: THE CASE OF CENTRAL AND SOUTHERN IRAQ

OTHMAN SAFAA AL-SHALASH

M.Sc. in Applied Statistics

Advisor: Asst. Prof. Dr.Cemal Murat ÖZKUT

July, 2020

To appropriately estimate the wind power output we need to invest more efforts in the determination of wind speed characteristics because such determination will play a valuable rule in the estimation approach. Following that, one of the most important distribution that has been implemented widely for modeling wind speed is Weibull distribution. This distribution is based mainly on statistical computations to model the wind speed in appropriate manner. Although it is used intensively in practical applications, the accuracy of it might not be completely optimal for modeling all wind regimes. Therefore, a variety of distributions that might play an alternative approach or in some cases a replacement for the Weibull distribution might model the wind speed in optimal way. In our research, we make use of other distributions which will be utilized as a different approaches from Weibull distribution for instance, Lognormal, Burr type XII, Generalized Extreme value, Gumbel, Gamma, Inverse Gaussian, and Rayleigh as alternatives to Weibull distribution. The main goal of our study is to define a suitable distribution, which provides the adequate realization of the inconstant behavior of the wind speed that might exist in different regimes. In the implementation section, several sets of data have been captured from Iraqi Meteorological Organization and Seismology – Ministry of Transportation for the previous ten years. To determine a convenient wind speed distribution, a variety of tests has been selected.

Keywords: The cumulative distribution function (cdf), the probability density function (pdf), Weibull distribution, Maximum likelihood, wind speed, model selection criteria.

ÖZET

RÜZGAR HIZI DAĞILIMLARININ PERFORMANSININ KARŞILAŞTIRILMASI: MERKEZ VE GÜNEY IRAK ÖRNEĞİ

OTHMAN SAFAA AL-SHALASH

Uygulamalı İstatistik, Yüksek Lisans Programı

Tez Danışmanı: Dr. Öğr. Üyesi. Cemal Murat ÖZKUT

Temmuz, 2020

Rüzgar enerjisi üretimini uygun şekilde tahmin etmek için rüzgar hızı karakteristiklerinin belirlenmesi için daha fazla çaba harcamamız gerekir, çünkü bu tür bir tespit tahmin yaklaşımında önemli bir rol oynamaktadır. Bunu takiben, rüzgar hızının modellenmesi için yaygın olarak uygulanan en önemli dağılımlardan biri Weibull dağılımıdır. Bu dağılım esas olarak rüzgar hızını uygun şekilde modellemek için istatistiksel hesaplamalara dayanmaktadır. Pratik uygulamalarda yoğun olarak kullanılmasına rağmen, doğruluğu tüm rüzgar rejimlerini modellemek için tamamen uygun olmayabilir. Bu nedenle, alternatif bir yaklaşım olabilecek veya bazı durumlarda Weibull dağılımının yerini alabilecek çeşitli dağılımlar, rüzgar hızını en uygun şekilde modelleyebilir. Araştırmamızda, Weibull dağılımından farklı bir yaklaşım olarak kullanılacak diğer dağılımlardan yararlanıyoruz (Lognormal, Burr tip XII, Genelleştirilmiş Aşırı değer, Gumbel, Gama, Ters Gauss, Genelleştirilmiş Aşırı değer, ve Rayleigh). Çalışmamızın temel amacı, farklı rejimlerde var olabilen rüzgar hızının tutarsız davranışının yeterli şekilde gerçekleştirilmesini sağlayan Weibull dağılımına alternatif olarak uygun bir dağılım tanımlamaktır. Uygulama bölümünde geçtiğimiz on yıl boyunca Irak Meteoroloji Örgütü ve Sismoloji - Ulaştırma Bakanlığı'ndan çeşitli veriler toplandı. Uygun rüzgar hızı dağılımını belirlemek için çeşitli testler seçildi.

Anahtar Kelimeler: Weibull dağılımı, kümülatif dağılım fonksiyonu (cdf), olasılık yoğunluk fonksiyonu (pdf), Maksimum olasılık, rüzgar hızı, model seçim kriterleri.

Dedicated to….

The candle that burns to illuminate my path, my father.

To the source of tenderness and the maker of men... My mother.

To those whose heart beats for me ... my brothers and sisters.

To my dear soul partner… my wife.

To my children ... my daughter Rama and my son Muhammad.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to Dr.Cemal Murat ÖZKUT for his guidance and insight throughout the research.

I would like to thank Prof. Gözde Yazgı TÜTÜNCÜ for her effort, energy, and time for supervising me diligently during my master's program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF CHARTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The energy requirements in our current study have been increased accordingly. This rise is strongly connected with the evolution of the technologies as well as the continuous increase in the population. Through the decades, we used a traditional approach to fulfil the energy requirements, which is based entirely of fossil fuels.

However, due to the environment pollution and climate change that has been arisen from the extensive usage of fossil fuels see Bilir, L., Devrim, Y. and Imir, M. (2015), Hepbasli, A. and Ozgener, O. (2004), based on that, many countries in the world intentionally invested in the renewable energy sources of the wind due to the fact this energy is considered a sustainable, clean, and cheap see Gani et al. (2016).

All around the world the wind energy has become the center of attention as the suitable renewable energy see Alavi, O., Mohammadi, K. and Mostafaeipour, A. (2016). For instance, multiple countries started already to produce electricity like, India, United States, Spain, Denmark, and Germany used the renewable source of energy, such us wind, to generate the electricity in clean way see Ahmed, A. S. (2010), and Bilir, L., Devrim, Y. and Imir, M. (2015).

The most important key factors for the wind energy system are represented by choosing the location after lengthy examination, so that we can determine exactly where the facility might be installed. The second factor is the understanding of the wind speed attribute Akgul, F., Arslan, T. and Senoglu, B. (2016), Ozay, C., and Celiktas. M. S. (2016).

Statistical distributions are essential to define the characteristics of the inconstant behavior of the wind speed. Regarding this, the modeling of the wind speed is based entirely in multiple scenarist on the Weibull distribution, see Bilir, L., Devrim,

Y. and Imir, M. (2015), Jowder, F. (2009), Costa Rocha .et al. (2012), Khahro et al. (2014), Bagiorgas et al. (2015), and Akdağ, S. and Güler, Ö. (2018).

Consequently, the first aim of our research will contrast the different distributions such as Rayleigh, Lognormal, Gumbel, Inverse Gaussian, Burr type XII, Generalized Extreme value, Gamma and Weibull in which we can determine the most promising one of those distributions in modelling wind speed.

We are encouraged to choose these distributions because they are applicable alternatives for the Weibull distribution, which has been used extremely in wind speed studies. Additionally, these distributions have been preferred in our study because they can model the datasets by taking into account the skewness and heavy tail.

The second aim that we have focused on was taking eight different stations in Iraq. These locations are distributed closely from each other where the wind speed performance is most attractive. More specifically, we acquired a multiple datasets for multiple locations and then we will try to study the attribute of each dataset to determine the wind speed characteristics correctly.

It is clear that Weibull distribution used to interpret and understand the behavior and characteristics of wind speeds that has been used widely, due to its properties that can be comprehended easily.

Based on that, the accuracy of each modeling distribution which is related more or less to the nature or the location of the wind regimes see Akgul, F., Arslan, T. and Senoglu, B. (2016), Gugliani, et al. (2017), Kantar, Y. and Usta, I. (2015), Ouarda, T. B., Charron, C., and Chebana, F. (2016). However, the aforementioned modeling systems have relative limitations when modeling the wind speed is faced with calm or extreme wind motions.

Furthermore, another concern of modeling the wind speed is associated with kurtosis as well as skewness see Chang, T. (2011), and Arslan, T., Acitas, S. and Senoglu, B. (2017). Hence, a wide range of the distributions have been implemented to realize the attribute of the wind speed for example, Beta, Nakagami, Gamma, Inverse Gaussian, generalized Gamma, Inverse Weibull (IW) , lognormal, Gumbel, Kappa, truncated normal, Wakaby, log-logistic, Rayleigh, and Inverse Gamma (IG).

The remaining sections of our research will be sorted in the following manners. In chapter 2, we made literatures review. In chapter 3, we will touch briefly the equations of the multiple distributions. In chapter 4, we evaluate our model by considering multiple criteria. In chapter 5, we will present the case study and examine the data of the wind speed that will be studied in our research. In chapter 6, results and discussions of our research are explained thoroughly. Lastly, we present our conclusion.

CHAPTER 2: LITRATURE REVIEW

Other authors examined other literatures that are related to the wind energy and wind speed characteristics for example Mahmood, F., Resen, A. and Khamees, A., (2020), who used Weibull distribution as a model to analyze the wind speed in AL Salman station, Iraq.

In addition, Hassoon A., (2013), was concerned with the assessment of wind energy in north of Iraq. They test five different locations and these data analyzed by using Weibull distribution.

Subsequently, Shu, Z., Li, Q. and Chan, P. (2015), were relied mainly on defining wind speed behavior on the Weibull distribution in Hong Kong. They were able through investigation and research in five different locations within six years by using Weibull model to discover that the annual average parameter scale is between (2.85 - 10.19) and the annual rate for shape parameter was between (1.65-1.99).

In the same context, Kidmo et al. (2015), in the Garoua region, Cameron. The authors used a Weibull model to explain the characteristics of wind speed. In order to predict Weibull parameters, they tested several methods such as Empirical Method, Energy Pattern Factor method, Moment Method Graphical Method, Modified Maximum Likelihood Method (MMLE) and Maximum Likelihood Method (MLE). They examined these predictions for Weibull parameters by utilizing several criteria's such as (coefficient of correlation, Kolmogorov-Smirnov (KS), root mean square error (RMSE), and chi-square) tests. The authors realized that the best prediction was with the Moment Method and Energy Pattern Factor method.

Furthermore, Shittu, O. and Adepoju, K. (2014), tried to utilize an Exponentiated Weibull distribution in the area located in South Western Nigeria as an alternative to Weibull distribution, which is considered as the most common distributions used in the interpretation of wind speed performance. The authors compared the performance of the two distributions systematically.

Hence, the values that have been shown from the use of two selection criterion such as likelihood function and Akaike information criterion (AIC), where the best fit for the data of wind speed. The criteria values for Exponentiated Weibull distribution were always lower than Weibull values over the Year, except one month only. This result encouraged the authors to consider Exponentiated Weibull distribution as a successful alternative to Weibull distribution.

From another point of view, the authors Zaharim et al. (2009) used a group of methods, which enables them to assess the wind speed data in comparative manner. This assessment took place in the Engineering Faculty, University Kebangsaan Malaysia. Additionally, they utilized the two-parameter Weibull distribution as well as lognormal distribution to model the data of wind speed. The results showed them that the two-parameter Weibull distribution is more reliable than Lognormal distribution.

On the other hands, the authors Carta, J., Ramírez, P. and Velázquez, S., (2009), and Morgan et al. (2011), have tried to sort out the modeling performance of wind speed by using a set of distributions. They grasped a very significant realization that the Weibull distribution might not be useful in all wind regimes. Following that, they determined a unique factor that need to be adhered to reduce the expected amount of errors in the wind estimation.

Brano et al. (2011) examined the speed characteristics in Palermo, Italy. Such examination was possible by the utilization of the following distributions: Rayleigh, Weibull, Gamma, Inverse Gaussian, Person (Type 5), Burr Type XII and Lognormal distributions. The authors demonstrated that one type of distribution, which is called Burr type XII distribution, was suitable for this area.

Further research was conducted by Amaya-Martínez, P., Saavedra-Montes, A. and Arango-Zuluaga, E. (2014), which took into account four common distributions, which are Lognormal, Weibull, Gamma, and Rayleigh. The authors carried out this study in the Aburra Valley, Colombia in five different stations to test the mentioned common distributions previously to realize which one of the distributions express the best fit for the data of the five sites of interest. The results showed that the Lognormal distribution got the first rank in three different locations, while the Weibull distribution came in second place as the best fit for the data. As for the Gamma distribution, it came third in only one location.

On the contrary, Mohammadi, K., Alavi, O. and McGowan, J. (2017), studied the wind speed characteristics in Ontario, Canada. They were constrained with one distribution, which was Birnbaum-Saunders (BS) distribution to compare its validity against varies distribution in this regimes. They concluded that BS distribution was the most convenient one.

Barcale, A., Carpinelli, G. and De Falco, P. (2017), utilized the inverse Burr distribution to estimate the wind speed characteristics. Moreover, Jung, C. and Schindler, D. (2017), took another approach by combining one component distribution and mixture distributions in which they can model the wind speed globally. They reached a conclusion that Weibull distribution has proven to be the best in handful regimes.

In addition to that, Arslan, T., Acitas, S. and Senoglu, B. (2017), used power Lindley PL distributions and Generalized Lindley GL. They emphasized that GL distribution is best fitting for test purpose and PL distribution was chosen to classify the power density error.

2.1 Parameters estimation method

There are multiple estimation methods to predict the parameters of distributions such as Empirical Method, Moment Method, least square method, Graphical Method, weighted least square method, Modified Maximum Likelihood Method (MMLE) and Maximum Likelihood Method (MLE). Following that we will explain each of these methods in simple manner in the following sections, for additional information see Pobocikova, I. and Sedliackova, Z. (2014), Teimouri, M., M, S. and Nadarajah, S., (2013), and Al-Fawzan, M., (2000).

Through the following methods, we will use an example, which will explain the Weibull distribution estimation parameters more effectively, the pdf and cdf for Weibull as below respectively:

$$f(x) = \frac{c}{\sigma}\left(\frac{x}{\sigma}\right)^{c-1} e^{-(\frac{x}{\sigma})^{c}} \qquad (2.1)$$

$$F(x) = 1 - e^{-\left(\frac{x^{c}}{\sigma^{c}}\right)} \qquad (2.2)$$

2.1.1 Graphical methods

These methods have been utilized due to the fact that they are simple to implement and fast to get the result. Although , it is simple and fast but it suffers a high percentage of error in the results. One of the most common graphical method is Weibull probability plotting. Weibull palpability plotting is based mainly on the logarithmic transformation. The equation (2.2) will be transformed to logarithmic function, which seen below:

$$\ln\left[-\ln\left(1 - F(x)\right)\right] = cln(x) - cln(\sigma) \qquad (2.3)$$

To determine the Weibull scale and shape parameters we will draw $ln(x)$ versus $\ln\left[-\ln\left(1 - F(x)\right)\right]$ so that we can acquired the straight line in the graph. Thus, each zero presence on the graph will be ignored to get line of the best fit. The c is defined as the slope of the line and the value of the term$(-cln(k))$ is defined as the y-intercept.

2.1.2 Least Squares Method

This method is another technique, which has been utilized to estimate several problem in the fields of engineering and mathematics. However, this method is not pure to estimate the parameters. To estimate the weibull parameters we will use the LSM and that is demonstrate in the following equations.

To convert equation (2.2) to linear formula, it will undergo a two logarithmic computation, and then we will get:

$$\ln\left[-\ln\left(1-F(x)\right)\right] = cln(x) - cln(\sigma) \qquad (2.3)$$

where, $\ln[-\ln(1-\mathrm{F(x)}) = \mathrm{Y}$, and $\ln(x) = X$, let $B_1 = c$ and $B_0 = -cln(\sigma)$, then we can rewrite the equation (2.3) as follows:

$$Y = B_1 X + B_0 \qquad (2.4)$$

Now let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ will be defined as the order statistics of $X_1, X_2, \dots \dots .., X_n$ and let $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$be the observations, which are ordered. The mean rank is used to assess the values of the CDF function as in the following equation:

$$\hat{F}\left(x_{(i)}\right) = \frac{i}{n+1} \qquad (2.5)$$

where i is defined as the i^{th}smallest value of $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, i= 1, 2… n. The estimate $\hat{B}_1$ and $\hat{B}_0$ of the regression, parameters B_1 *and* B_0 minimize the equation as below:

$$Q(B_0, B_1) = \sum_{i=1}^{n} (Y_i - B_0 - B_1 \ln x_i)^2 \qquad (2.6)$$

The estimate $\hat{B}_0$ and $\hat{B}_1$ of the parameters B_0 *and* $B_1$are givern by:

$$\hat{B}_1 = \frac{n\sum_{i=1}^{n} lnx_i \ln[-\ln(1-\hat{F}_{(x_i)}))] - \sum_{i=1}^{n} lnx_i \sum_{i}^{n} \ln(-\ln(1-\hat{F}_{(x_i)}))]}{n\sum_{i=1}^{n} ln^2 x_{(i)} - \left(\sum_{i=1}^{n} lnx_{(i)}\right)^2} \qquad (2.7)$$

$$\hat{B}_0 = \frac{1}{n}\sum_{i=1}^{n} \ln[-\ln(1-\hat{F}_{(x_i)}))] - \hat{B}_1 \frac{1}{n}\sum_{i=1}^{n} \ln x_{(i)} \qquad (2.8)$$

where $\hat{B}_1 = \hat{c}$ and $\hat{\sigma} = \exp\left(\frac{\ln[-\ln\left(1-\hat{F}_{(x_i)}\right)-\hat{c}\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)}}{\hat{c}n}\right)$

2.1.3 Weighted least square method

This method is commonly applied in the field of estimated parameters because it is remarkably simple for the estimation. Additionally the calculation of the estimation can be obtained very easily by using closed-form formula. The estimate $\hat{B}_1$ and $\hat{B}_0$of the regression, parameters $B_1\ and\ B_0$ minimize the function:

$$Q(B_0, B_1) = \sum_{i=1}^{n} w_i (Y_i - B_0 - B_1 \ln x_i)^2 \qquad (2.9)$$

Let, w_i the weight factor, which is described in the equation as bellows:

$$w_i = \left[(1 - \hat{F}_{(x_i)})\ln(1 - \hat{F}_{(x_i)})\right]^2 \qquad (2.10)$$

The predict $\hat{B}_1$ and $\hat{B}_0$ of the parameters $B_1\ and\ B_0$ are givern by:

$$\hat{B}_1 = \frac{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i lnx_i \ln[-\ln(1 - \hat{F}_{(x_i)}))] - \sum_{i=1}^{n} w_i lnx_i \sum_{i}^{n} w_i \ln(-\ln(1 - \hat{F}_{(x_i)}))]}{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i ln^2 x_{(i)} - \left(\sum_{i=1}^{n} w_i lnx_{(i)}\right)^2} \qquad (2.11)$$

$$\hat{B}_0 = \frac{\sum_{i=1}^{n} w_i \ln[-\ln(1 - \hat{F}_{(x_i)}))] - \hat{B}_1 \frac{1}{n} \sum_{i=1}^{n} w_i \ln x_{(i)}}{\hat{B}_1 \sum_{i=1}^{n} w_i} \qquad (2.12)$$

where $\hat{B}_1 = \hat{c}$ and $\hat{\sigma} = \exp\left(\frac{\sum_{i=1}^{n} w_i \ln[-\ln\left(1-\hat{F}_{(x_i)}\right)-\hat{c}\sum_{i=1}^{n} w_i \ln x_{(i)}}{\hat{c}\sum_{i=1}^{n} w_i}\right)$

2.1.4 Method of Moments

This method is one of the most promising techniques that has been implemented widely to estimate the parameters. Let $x_1, x_2, \dots \dots .., x_n$ be a set of data, by using v_m, which is the mean, and sigma, which is the standard deviation of Weibull distribution to estimate parameters for moment method. To solve the moment method we can use the iteration techniques by the following equation:

$$\sigma = \frac{v_m}{\Gamma(1 + 1/c)} \qquad (2.13)$$

We can obtain the standard deviation sigma from the equation below:

$$sigma = \sigma[\Gamma\left(1+\frac{2}{c}\right) - \Gamma^2(1+1/c)]^2 \qquad (2.14)$$

where the gamma function is written as below:

$$\Gamma(x) = \int_0^\infty t^{x-1}\exp(-t)dt \qquad (2.15)$$

2.1.5 Modified Maximum likelihood Estimation Method (MMLE)

This method is restricted only for the data of wind speed, which is visible in the Weibull distribution format. (MMLE) uses numerical iteration to estimate the parameters such us for Weibull distribution. Furthermore, the shape and scale parameters can be estimated by the following equations:

$$c = \left[\frac{\sum_{i=1}^{n} x_i^c \ln(x_i) f(x_i)}{\sum_{i=1}^{n} x_i^c f(x_i)} - \frac{\sum_{i=1}^{n} \ln(x_i) f(x_i)}{f(x \geq 0)}\right]^{-1} \qquad (2.16)$$

$$\sigma = \left[\frac{\sum_{i=1}^{n} x_i^c f(x_i)}{f(x \geq 0)}\right]^{-1} \qquad (2.17)$$

Here, $f(x)$ is the probability density function when $x \geq 0$, $f(x_i)$ is Weibull frequency.

2.1.6 Empirical Method

This method is considered a special case of the moment method, where the Weibull parameters c and σ can be shown by the equations below:

$$c = (sigma/v_m)^{-1.089} \qquad (2.18)$$

$$\sigma = \frac{v_m}{\Gamma(1+1/c)} \qquad (2.19)$$

where, v_m the mean and sigma be the standard deviation of Weibull distribution.

2.1.7 Maximum likelihood Estimation method (MLE)

This method is one of most promising method, which will be used to maximize the likelihood function by taking into account several significant parameters.

Historically, Pierre-Simon Laplace, Carl Friedrich Gauss, Francis Ysidro Edgeworth and Thorvald N. Thiele were the first scientists who using the maximum likelihood estimation method, for more information see Edgeworth, F., (1908).

However, the use of maximum likelihood estimation increased widely between 1912 and 1920 when R. A. Fisher considered careful analysis of the maximum likelihood estimation see Pfanzagl, J. and Hamböker, R., (1994). It is worth noting, that the first person to represent the maximum likelihood estimation is the scientist R. A. Fisher in 1922, when he made many researches and developments on the way until he reached the current form see Aldrich, J. (1997).

It is worth to note that many authors recommended using this method (maximum likelihood estimation) to determine the parameters due to its distinctive characteristics which are represented through efficiency, consistency and asymptotic normality. Additionally, the maximum likelihood estimation is considered to be feasible when the sample size is not small see Pobocikova, I. and Sedliackova, Z., (2014).

Now, we are going to introduce this method briefly just to realize and understanding the process. Maximum likelihood estimation (MLE) is a procedure to predict parameters of a distribution, besides making test hypothesis for parameters, which are estimated. In addition, like any method, maximum likelihood estimation need two fundamental components. The first one is a mathematical model, which describe the distribution of the variables of the data set. The second one is a set of data that is obviously important for any statistical analysis.

We can define the parameters as an unknown quantity, which are belong to model, which describes the distribution in database. In general, the aim of maximum likelihood estimation (MLE) method is to consider the parameters of our model, which provides the best interpretation for the data. Consequently, in data analysis, three different steps are used in maximum likelihood estimation. Firstly, the parameters estimation for distributions. Secondly, making test hypothesis for those parameters. Lastly, on the same data that are used to estimate parameters, making comparison on two models.

In the estimation part, it is probably best to set an example for the estimation process to be clear and easy to understand. Suppose that the variable X_i in a data set is normally distributed, n is the number of observation and assume we tried to estimate the normal distribution parameters, which are Standard Deviation and Mean.

$$L(X_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X_i-\mu}{\sigma}\right)^2} \qquad (2.20)$$

where, μ is sign to Mean, σ is sign to the standard deviation of the distribution. Suppose that the observations (x_1,x_2,..........x_n) are independent so we can write the joint likelihood as follows:

$$f(x_1, x_2, \dots \dots \dots, x_n|\theta) = f(x_1|\theta)f(x_2|\theta) \dots \dots \dots \dots \dots f(x_n|\theta) \qquad (2.21)$$

where, θ is unknown parameters and $f(x_i|\theta)$ are a probability density functions according to ith of observations. It is normal to denote to $f(x_1, x_2, \dots \dots \dots, x_n|\theta)$ by L(X) so we have got:

$$L(x) = \prod_{i=1}^{n} L(X_i) \qquad (2.22)$$

Now, we have to maximize the function with respect to unknown parameters , μ and σ by taking the first derivative to equation respecting to unknown parameters and equate them to zero , such that :

$$\frac{\partial L(x)}{\partial \mu} = 0 , \frac{\partial L(x)}{\partial \sigma} = 0$$

Before going to take derivatives, we should use Logarithm function for maximum likelihood estimation method because of three reasons to deal with log function instead of deal with likelihood function directly. First, the log function is a monotonic increasing function. Second, log likelihood simplifies the comparison with derivative calculation of likelihood function because log likelihood is a series of sums while the likelihood is series of products**.** Third, likelihoods functions are often (but not always) quantities between 0 and 1. Hence, taking the products of a large number of fractions can be affected by marked rounding and truncation error, even with the most modern computers**.** So, to come back to our example after taking Logarithm to equation (2.22) we will get:

$$Log\big(L(x)\big) = \sum_{i=1}^{n} Log(L(x_i)) \qquad (2.23)$$

By substituting (.220) in (.223) we can get:

$$Log\big(L(x)\big) = \sum_{i=1}^{n} (\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{1}{2}\Big(\frac{x_i - \mu}{\sigma}\Big)^2)$$

$$Log\big(L(x)\big) = (\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\Big(\frac{x_i - \mu}{\sigma}\Big)^2) \qquad (2.24)$$

Finally, taking the derivative is the next step respecting to μ and σ and equate them to zero we will get:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (2.25)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}} \qquad (2.26)$$

where, $\hat{\mu}$ is The estimator of μ, $\hat{\sigma}$ is the estimator of σ.

It is worth noting that the solution to derivatives is almost very difficult or in other words, the solution of these derivatives requires a great effort to solve them. Interestingly, the result in the previous example showed the ease of finding to statistical expressions $\hat{\mu}$ and $\hat{\sigma}$ which are almost rare and unusual.

As a result, many numerical methods and techniques have been used to find Maximum likelihood estimators, and these popular and widely used methods such as, Bisection method, Newton-Raphson method, and secant method.

CHAPTER 3: WIND SPEED DISTRIBUTIONS

In this part from our research, we will talk briefly about the probability density functions (pdf) and cumulative distribution function (cdf) for each distributions, which are used in this study. It worth to note that statistical distributions has a significance by interpreting the wind speed in appropriate way.

3.1 Weibull distribution

Weibull distribution (WD) is used to describe various kinds of observed failures of data. In addition, these models have several use cases in reliability and survival data, likewise in reliability engineering see (Pham, 2007).

It is worth noting that Weibull distribution have a lot of applications that we will mention some of them are as follows: wind speed data analysis see Al-Hasan, M. and Nigmatullin, R., (2003), Earthquake magnitude see Huillet, T. and Raynaud, H., (1999), Survival data see Carroll, K., (2003), Environment radioactivity see Dahm, H., Niemeyer, J. and Schröder, D. (2002), and in nature see Fleming, R., (2001).

As we mentioned above, Weibull distribution has been considered the prevalent statistical distribution in which it models the speed of the wind because it was proven in many cases in which it was realized to be optimal in the nature, see Acker et al. (2007), Ahmed Shata A.S., Hanitsch R. (2006), and Akpinar EK, Akpinar S. (2005). The pdf as well as cdf written in the following way:

$$f(x) = \frac{c}{\sigma}\ (\frac{x}{\sigma})^{c-1}\ e^{-(\frac{x}{\sigma})^{c}} \qquad (3.1)$$

$$F(x) = 1 - e^{-\left(\frac{x^{c}}{\sigma^{c}}\right)} \qquad (3.2)$$

where, x >0 and σ >0. c will interpreted as shape. σ will interpreted as scale. To maximize the joint likelihood function we utilized the ML estimation method to estimate unknown of Weibull distribution parameters. ln L function is considered as below:

$$\ln L = nlnc - ncln\sigma + (c-1)\sum_{i=1}^{n} lnx_i - \sigma^{-c}\sum_{i=1}^{n}(x_i^c) \qquad (3.3)$$

Then, we have to take derivatives of ln L function for c as well as for σ. Afterwards, the equations will be equalized to zero. As a result, the ML has been acquired. $\hat{c}$ will be acquired from the next equations, which are solved iteratively:

$$\hat{c} = \frac{n}{c} + \sum_{i=1}^{n} lnx_i - \frac{n\sum_{i=1}^{n} x_i^c lnx_i}{\sum_{i=1}^{n} x_i^c} = 0 \qquad (3.4)$$

After this, by incorporating $\hat{c}$ into next equation we obtained σ hat as follows:

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n} x_i^{\hat{c}} \qquad (3.5)$$

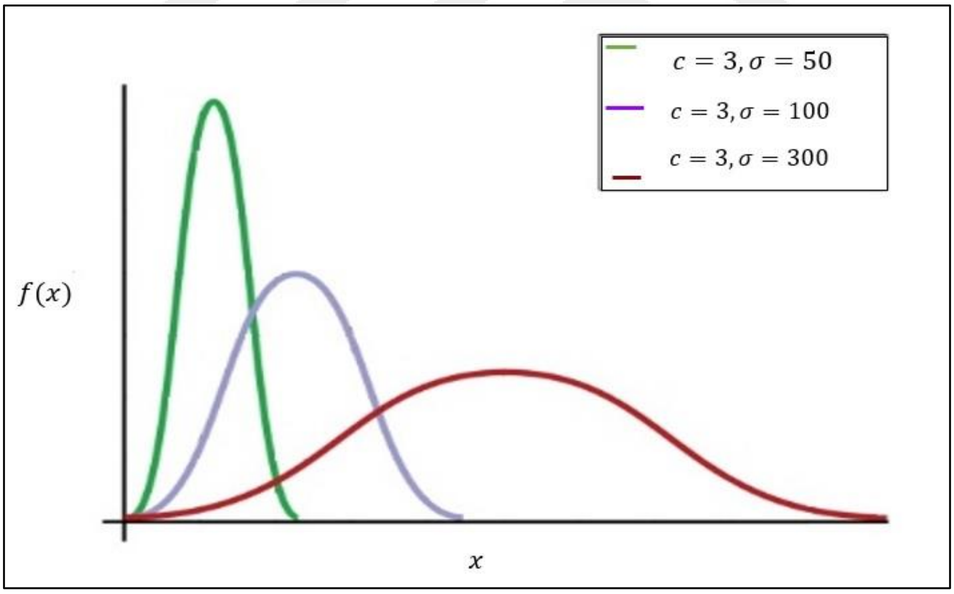

Figure 1**:** Performance the pdf s of Weibull distribution 1.

By examining figure 1, we can see the performance of distribution when scale parameter getting higher and the shape parameter kept as a constant. Following that, it will lead that the distribution gets stretched out to the right and its height decreases, while maintaining its location. When the scale parameter went down and the shape parameter kept constant, it will derive that, the height of the distribution to be increase and pushes it to the left.

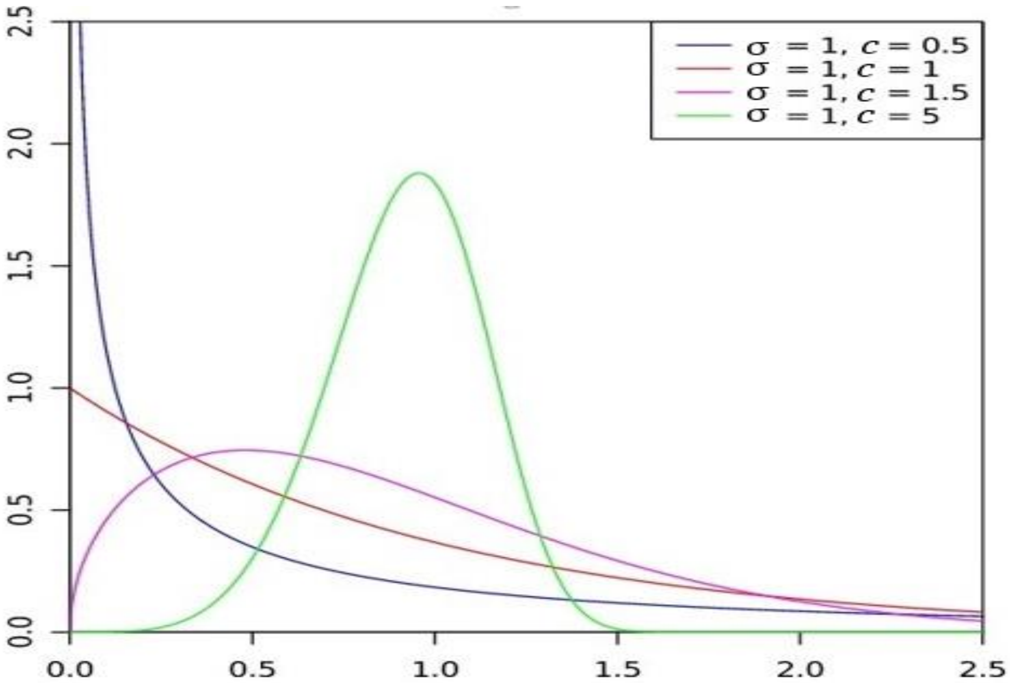

Figure 2: Performance the pdf s of Weibull distribution 2.

By looking into figure 2, we be able to realize that the shape parameter is effected clearly on the attribute of Weibull distribution when the values of shape parameters are chosen differently. Moreover, in some literatures the shape parameter is called as slope parameter.

In fact, new distribution equations can be obtained by simply changing some values of the shape parameters, for example, exponential distribution is considered as a special case of Weibull distribution, which can be acquired when the shape parameter is equal to 1. In addition, Rayleigh distribution is also considered as a special case from Weibull distribution when the value of the shape parameter equal to 2. Furthermore, when the values of the shape parameter are between 3 and 4 will be approximately symmetric.

3.2 Rayleigh distribution

Rayleigh distribution is one of the most popular distributions used in statistical analysis, especially in the analysis of positive skewed data. In addition to this, Rayleigh distribution has become widely used in many applications, particularly oceanography applications and communication theory applications. It has been used to describe the instantaneous peak power of signals. However, it appears historically, Lord Rayleigh used the Rayleigh distribution initially in the areas of acoustics and optics.

The widespread use of Rayleigh encouraged many physicists and engineers to pay attention to this distribution and use it in modeling wave propagation, artificial radar images, radiation, etc.

In the same context, Rayleigh has been used extensively to model the characteristics that define the speed of the wind see Bidaoui, et al. (2019), and Akgul, F., Arslan, T. and Senoglu, B. (2016). The Rayleigh distribution, which is a distribution of continuous probability density function and it is special case of Weibull distribution with a fixed shape parameter value equal to 2. It is named after the English Lord Rayleigh. The Probability density function and the Cumulative distribution function for Rayleigh distribution are demonstrated as follows respectively:

$$f(x) = \left(\frac{x}{\sigma^2}\right) e^{-\frac{x^2}{2\sigma^2}} \qquad (3.6)$$

$$F(x) = 1 - e^{-\frac{x^2}{2\sigma^2}} \qquad (3.7)$$

where, x >0 and σ >0. Here, σ is scale parameter. To estimate the unknown parameter, ln L function is considered as below:

$$\ln L = -2nln\sigma + \sum_{i=1}^{n} lnx_i - \sum_{i=1}^{n} \left(\frac{{x_i}^2}{\sigma^2}\right) \qquad (3.8)$$

Then, we have to take derivatives of ln L function for σ. Afterwards, the equation will be equalized to zero. As a result, the ML equation has been acquired. $\hat{\sigma}$ will be acquired from the next equation:

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2} \qquad (3.9)$$

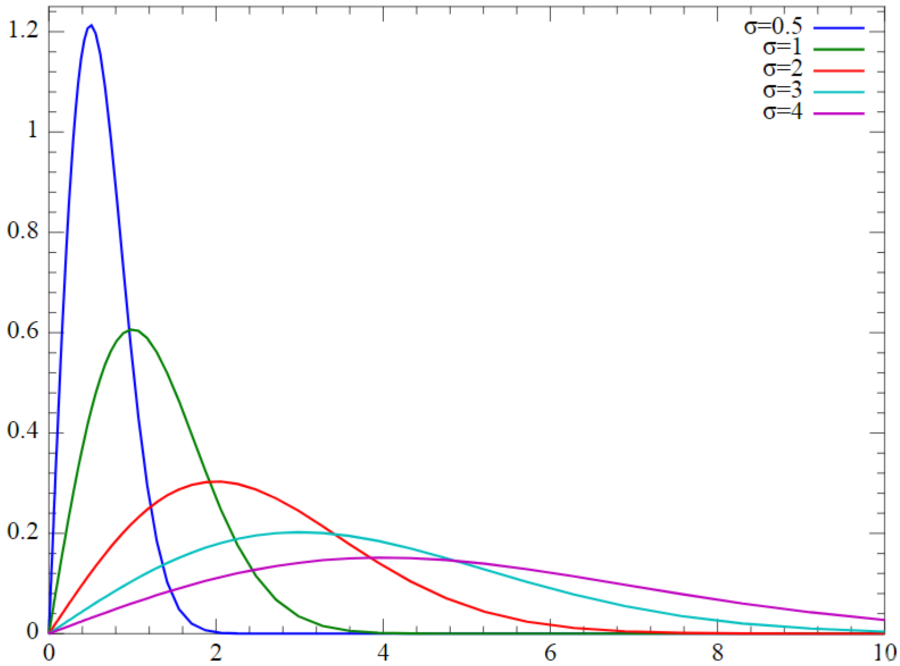

Figure 3: Performance the pdf s of Rayleigh distribution.

In figure 3, we realize the performance of Rayleigh distribution when scale parameter getting higher, the distribution gets stretched out to the right and its height will be decreased.

3.4 Lognormal Distribution

Lognormal distribution is known in some fields as the Galton distribution, and it is one of the continuous distributions used to model continuous random variables that are often greater or equal to zero.

Lognormal distribution has many applications in fields that are almost extensive and we will mention here some of them. Lognormal distribution was found useful in applications of environmental sciences, applications of chemicals, medicine, and economics. In the same context, the lognormal distribution was a respectable distribution in financial applications see Black, F., Scholes, M. (1973).

The lognormal distribution is distinguished by its use in the modeling of data that has been distributed normally, regardless of whether it is a little or more skewed see Ginos, (2009).

On the other hands, lognormal distribution has been considered the prevalent statistical distribution in which it models the speed of the wind see Ivana, P., Zuzana, S. and Mária, M. (2017), Ginos, (2009). Furthermore, the pdf and cdf for lognormal distribution written in the following way:

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-(\ln(x)-\mu)^2/2\sigma^2)} \qquad (3.10)$$

$$F(x) = \Phi\left(\frac{\ln(x) - m}{\sigma}\right) \qquad (3.11)$$

Here μ and σ mean and standard deviation, while Φ is the standard normal (standard Lognormal) distribution cdf. The ln L function is written as demonstrated below:

$$\ln L = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\frac{\ln(x_i)^2}{2\sigma^2} + \sum_{i=1}^{n}\frac{\ln(x_i)\mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2} \qquad (3.12)$$

After that, derivatives are taken for ln L function for μ as well as σ. Afterwards, the equations will be equalized to zero. As a result, the likelihoods equations can been written as follows:

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\ln(x_i)}{n} \qquad (3.13)$$

$$\hat{\sigma} = \frac{\sum_{i=1}^{n}\left(\ln(x_i) - \sum_{i=1}^{n}\frac{\ln(x_i)}{n}\right)^2}{n} \qquad (3.14)$$

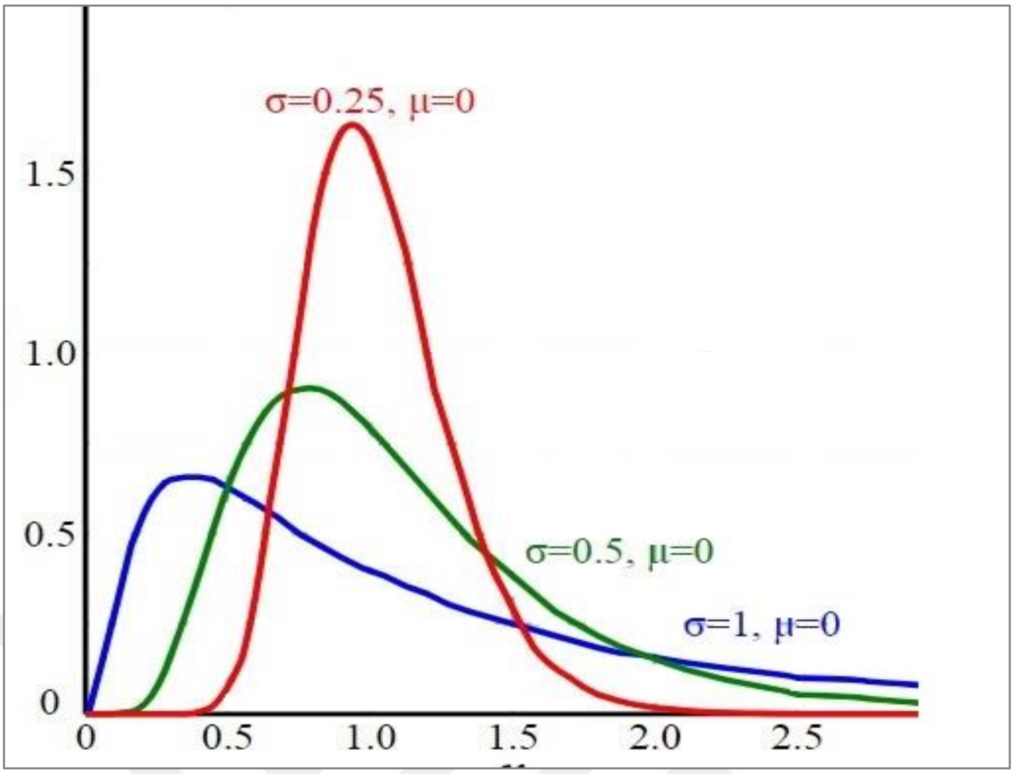

Figure 4: Performance the pdf s of Liognormal distribution.

In figure 4, we observe that the performance of probability density functions will become almost normal when the Mean parameter is equal to 0 and standard deviation is going down. While the performance of probability density function gets stretched out to the right and its height will be decreased when the standard deviation be bigger.

3.5 Inverse Gaussian distribution

Inverse Gaussian distribution is one of the distributions belonging to the family of two-parameter continuous distributions such that μ is Mean parameter and λ is considered as the shape parameter. Inverse Gaussian distribution is also known by the name of Wald distribution. For an optimal use of the inverse Gaussian distribution, there are some conditions that must be met by the data. In particularly, the data must be non-negative and positively skewed.

Furthermore, Inverse Gaussian distribution (IG) utilized as an alternative to weibull distribution especially in recent decades for modelling wind speed data see Philippe D., (2015). The pdf as well as cdf written in the following way:

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}} \qquad (3.15)$$

$$F(x) = \int_0^v \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}} dx \qquad (3.16)$$

Here, μ and λ mean and shape respectively. To find the ML estimates of parameters of Inverse Gaussian, which are unknown, we derive log function to maximize likelihood function as listed below:

$$\ln L = \frac{n}{2}\ln\left(\frac{\lambda}{2\pi}\right) - \frac{3}{2}\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i} \qquad (3.17)$$

Next, derivatives are taken for ln L function for μ as well as λ . Afterwards, the equations will be equalized to zero. As a result, ML equations has been obtained in the following way:

$$\hat{\mu} = \frac{n\lambda}{\mu^3}\left(\frac{\sum_{i=1}^{n} x_i}{n} - \mu\right) \qquad (3.18)$$

$$\hat{\lambda} = \frac{n}{2}\frac{1}{\lambda} - \frac{1}{2\mu^2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{x_i} \qquad (3.19)$$

In figure 5, we observe that the performance of probability density distributions when the parameter μ be a constant and standard deviation parameter λ is increased the pdf gets stretched out to the right Additionally, the effective appears that the height of pdf is increase when the parameter μ be a constant and standard deviation parameter λ is decrease.

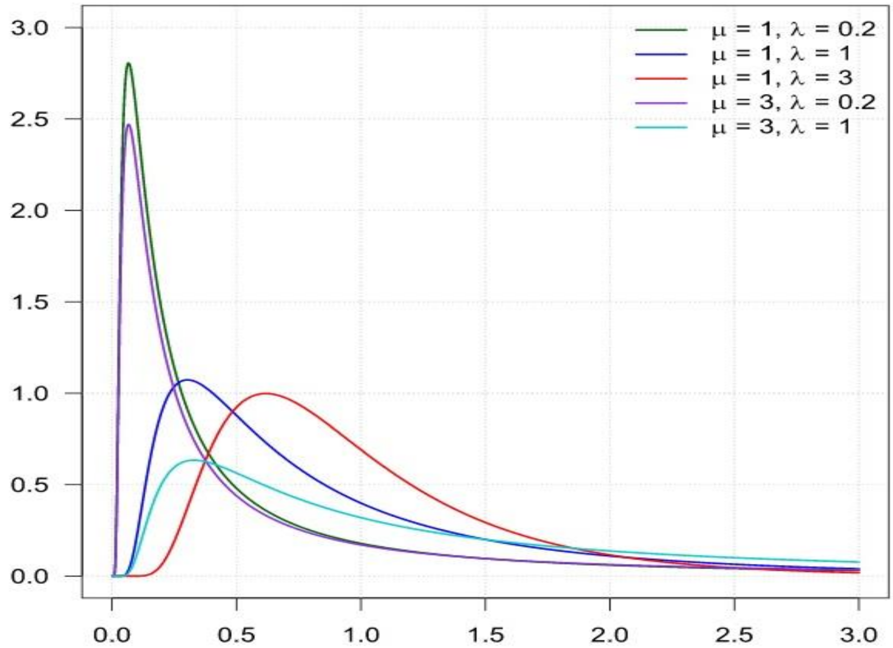

Figure 5: Performance the pdf s of IG distribution.

3.6 Generalized extreme value distribution

Generalized extreme value distribution (GEV) is a distribution that belongs to the continuous probability distributions family. It has been noted that the generalized extreme value distribution (GEV) is called Fisher - Tippett distribution in some books and articles, according to Ronald Fisher and L. C. C. Tippett. Those authors were able to obtain three different forms of generalized extreme value distribution (GEV).

For generalized extreme value distribution (GEV), there are wide applications in different fields, including for example, its uses in hydrology - where generalized extreme value distribution is utilized to illustrate the daily rainfall - . On the other hand, GEV is used for financial transfer and insurance modeling, as well as it is considered one of the best methods used to analyze financial risks through special criteria for example the value at risk, see Moscadelli, M., (2004), and Guégan, D., Hassani, B.K. (2014).

Broadly speaking, Generalized extreme value distributions have been widely used for fitting the distributions of extreme wind speed see Cheng, E. and Yeung, C.

(2002), and Sarkar et al. (2019). The pdf as well as cdf for GEV obtained in the following way respectively:

$$f(x) = \frac{1}{\sigma}\left[1 + c\left(\frac{x-\mu}{\sigma}\right)\right]^{-1-1/c} e^{-[1+c(x-\mu/\sigma)]^{-1/c}}, c \neq 0, \sigma \neq 0\,,$$

$$-\infty < \mu < \infty \quad 1 + c\left(\frac{x-\mu}{\sigma}\right) > 0, \qquad (3.20)$$

$$F(x) = e^{-[1+c(x-\mu/\sigma)]^{-1/c}} c \neq 0, \sigma \neq 0\,,$$

$$-\infty < \mu < \infty \quad 1 + c\left(\frac{x-\mu}{\sigma}\right) > 0 \qquad (3.21)$$

Here, c is the shape, μ is the location and σ is the scale parameters. The logarithm likelihood (ln L) function is written as follows:

$$\ln L = -nln\sigma - \frac{c+1}{c}\sum_{i=1}^{n} lnz_i - \sum_{i=1}^{n} {z_i}^{-1/c} \qquad (3.22)$$

where, $z_i = 1 + c(x_i - \mu/\sigma)$. Then, by taking derivatives for ln L function with respect to unknown parameters c, μ and σ and equating them to zero, then we obtained non-linear system equations as follows:

$$-\frac{c+1}{c\sigma}\sum_{i=1}^{n}(x_i-\mu){z_i}^{-1} - \frac{1}{c^2}\sum_{i=1}^{n} lnz_i {z_i}^{-\frac{1}{c}} + \frac{1}{c\sigma}\sum_{i=1}^{n}(x_i-\mu){z_i}^{-1-\frac{1}{c}} = 0 \qquad (3.23)$$

$$\frac{c+1}{\sigma}\sum_{i=1}^{n}{z_i}^{-1} - \frac{1}{\sigma}\sum_{i=1}^{n}{z_i}^{-1-1/c} = 0 \qquad (3.24)$$

$$-\frac{1}{\sigma} + \frac{c+1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu){z_i}^{-1} - \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu){z_i}^{-1-\frac{1}{c}} = 0 \qquad (3.25)$$

Finally, to solve the preceding non-linear system equations we used the iterative method simultaneously.

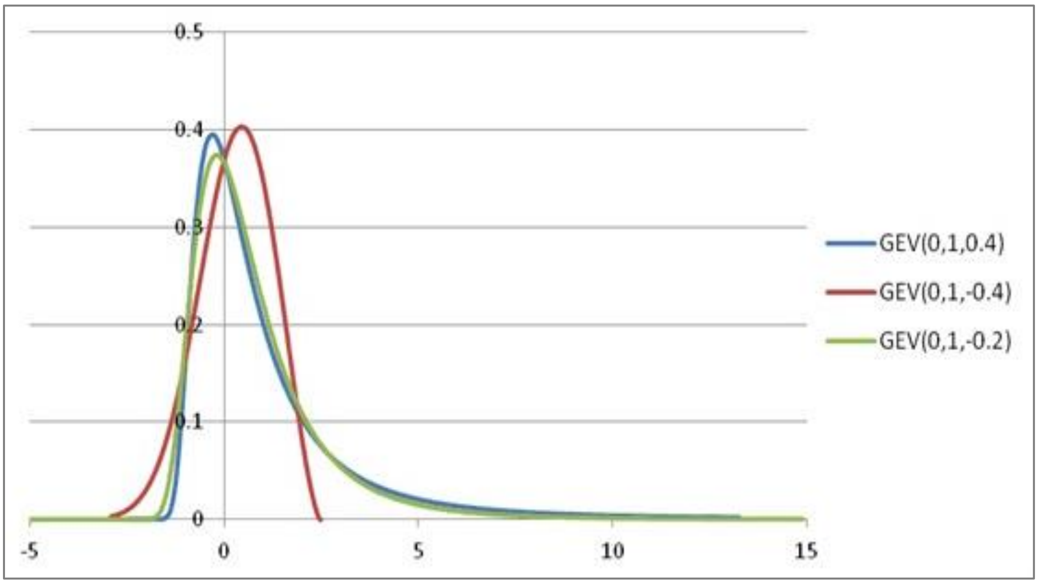

Figure 6: Performance the pdf s of GEV distribution.

In Figure 6, it represents the performance of probability density functions when location parameter = zero, scale parameter = one, and shape parameter takes a different values (0.4,-0.4, -0.2).

3.7 Gamma distribution

In statistics and probability theory, gamma distribution is a two-parameter distribution that belongs to continuous probability distributions and which has special cases such as Erlang distribution, exponential distribution, and CHI square distribution. Based on that, it has gained great importance and it is related to the aforementioned distributions and Normal distributions.

The Gamma distribution with two parameters (shape and scale) is one of the widely used distributions, especially in economics as well as in many other application fields. For example, modeling waiting times. In another example of applications, gamma distribution has been utilized in life tests, where the waiting time to death is a random variable (Hogg, 1978).

On the other hand, gamma distribution was also used in oncology, where it was found that the age distribution has been already followed by gamma distribution as cases of cancer, Belikov, A., (2017). In neuroscience, the use of gamma distribution was famous for its use in this field to describe the distribution of inter-spike intervals see Wright et al. (2014).

Following that, the gamma distribution is the most widely utilized distribution for modelling wind speed encountered in nature see Morgan et al.(2011), and Ivana, P., Zuzana, S. and Mária, M. (2017). The pdf and the cdf for Gamma distribution are obtained as follows respectively:

$$f(x) = \frac{1}{\Gamma(c)\sigma^c} x^{c-1} e^{-x/\sigma} \qquad (3.26)$$

$$F(x) = \frac{1}{\Gamma(c)\sigma^c} \gamma\left(c, \frac{x}{\sigma}\right) \qquad (3.27)$$

where, $x > 0, c > 0, and\ \sigma > 0$, c will be the shape parameter ,σ will be the scale parameter. $\Gamma(.)$ will be considered as gamma function . $\gamma(.)$ will be considered as the lower incomplete gamma function . The Log likelihood function (ln L) can be written in the following way:

$$\ln L = -nln\Gamma(c) - ncln\sigma + (c-1)\sum_{i=1}^{n} lnx_i - \sum_{i=1}^{n}\left(\frac{x_i}{\sigma}\right) \qquad (3.28)$$

Then, by taking derivatives of ln L function with respect to c and σ and equating them to zero, the $\hat{c}$ is obtained by solving this equation:

$$\psi(c) + ln\left(\frac{\sum_{i=1}^{n} x_i}{nc}\right) - \frac{1}{n}\sum_{i=1}^{n} lnx_i = 0 \qquad (3.29)$$

Now, to find the ML estimates of c we will utilize an iterative method. $\psi(c)$will be considered as digamma, which is designated for the Gamma function to be the logarithmic derivative, and c most be positive $c > 0$.the second estimated parameter $\hat{\sigma}$ is found by inserting $\hat{c}$ into next equation:

$$\hat{\sigma} = \frac{\sum_{i=1}^{n} x_i}{n\hat{c}} \tag{3.30}$$

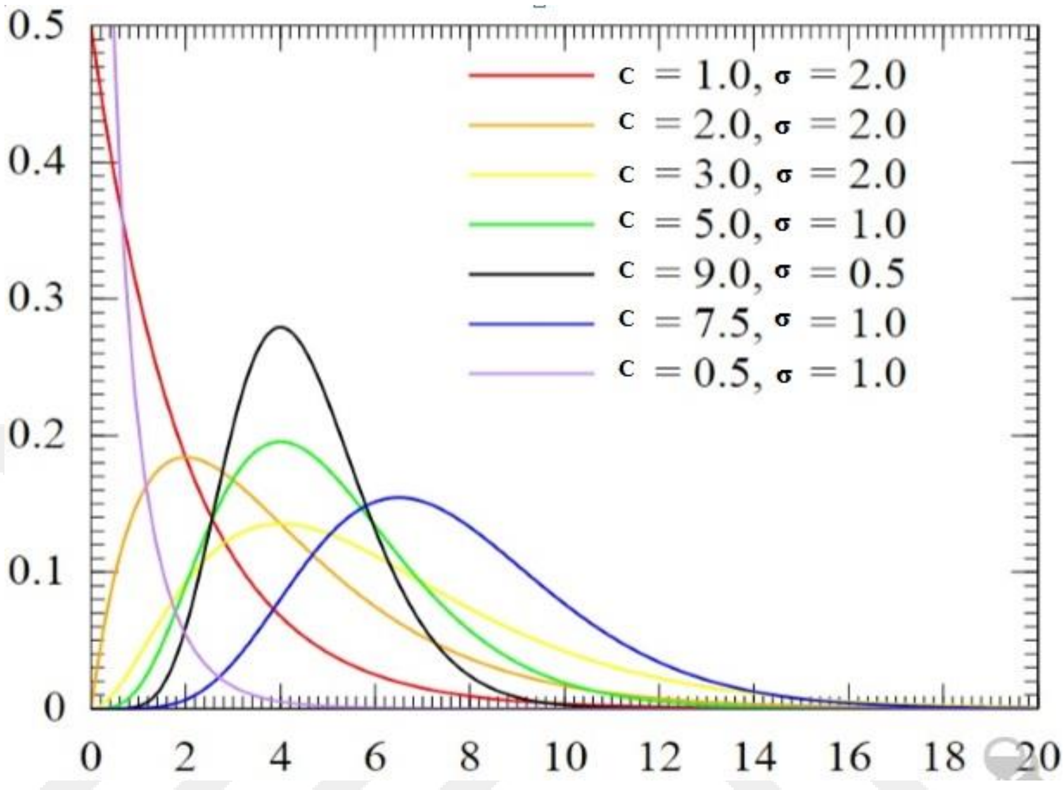

Figure 7: Performance the pdf s of Gamma distribution.

In figure 7, we can see the effective on probability density functions when shape and scale parameters take a different values. By examining figure (10), we can see the performance of distribution when shape parameter getting higher and the scale parameter kept as a constant. Following that, it will lead that the distribution gets stretched out to the right and its height decreases.

3.8 Gumbel distribution

Another alternative are used to describe the wind speed characteristics is Gumbel distribution. Initially, Gumbel distribution has utilized to describe one of the minimum distribution or maximum distribution of several different samples of distributions and all of this is implied by the theory of probabilities and statistics.

It is worth to note that The Gumbel distribution is a special case of the generalized extreme value distribution and there are three different types of Extreme Value Distribution, firstly, Gumbel distribution has been identified as most common Extreme value distribution in modeling the pollution degree, which might entail a valid result for the coming future.

Secondly, Fréchet Distribution has been utilized to model a wide range of applications in which it determines the maximum values in each form of data. Thirdly, the Weibull distribution is a very rich distribution due to the fact that it is used widely in the real world.

We clearly realized that Gumbel distribution has many practical applications. More specifically, the Extreme value represents the best option in analyzing the variables of the maximum values of the amount of daily, monthly or yearly rains and the volume of river discharge.

At the same context, the Gumbel distribution is named after Emil Julius Gumbel (1891–1966), and that is through the article he published that explained and described the distribution see Gumbel E.J. (1941). It is widely used as an extreme value distribution, which is used in modeling extreme wind speeds see Xiao et al. (2006), and Kang, D., Ko, K. and Huh, J., (2015).

The pdf as well as the cdf for Gumbel distribution are obtained as follows respectively:

$$f(x) = \frac{1}{\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)} e^{-e^{-\left(\frac{x-\mu}{\sigma}\right)}}, -\infty < x < \infty, , -\infty < \mu < \infty, \sigma > 0 \qquad (3.31)$$

$$F(x) = e^{-e^{-\left(\frac{x-\mu}{\sigma}\right)}}, -\infty < x < \infty, , -\infty < \mu < \infty, \sigma > 0 \qquad (3.32)$$

Here, μ and σ are the location and scale parameters respectively. To maximize the joint likelihood function we utilized the ML estimation method to estimate unknown parameters of Gumbel distribution. ln L function is showed as below:

$$\ln L = -n ln\sigma - \sum_{i=1}^{n} r_i - \sum_{i=1}^{n} (e^{-r_i}) \tag{3.33}$$

Here, $r_i = ((x_i - \mu)/\sigma$ .Then, we have to take derivatives of ln L function for our parameters μ as well as σ. Afterwards, the equations will be equalized to zero. As a result, ML equations has been obtained in the following way:

$$\frac{n}{\sigma} - \frac{1}{\sigma} \sum_{i=1}^{n} e^{-r_i} = 0 \tag{3.34}$$

$$-\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} r_i - \frac{1}{\sigma} \sum_{i=1}^{n} r_i e^{-r_i} = 0 \tag{3.35}$$

Then, equations (3.43) and (3.35) are solved at the same time iteratively by using numerical methods.

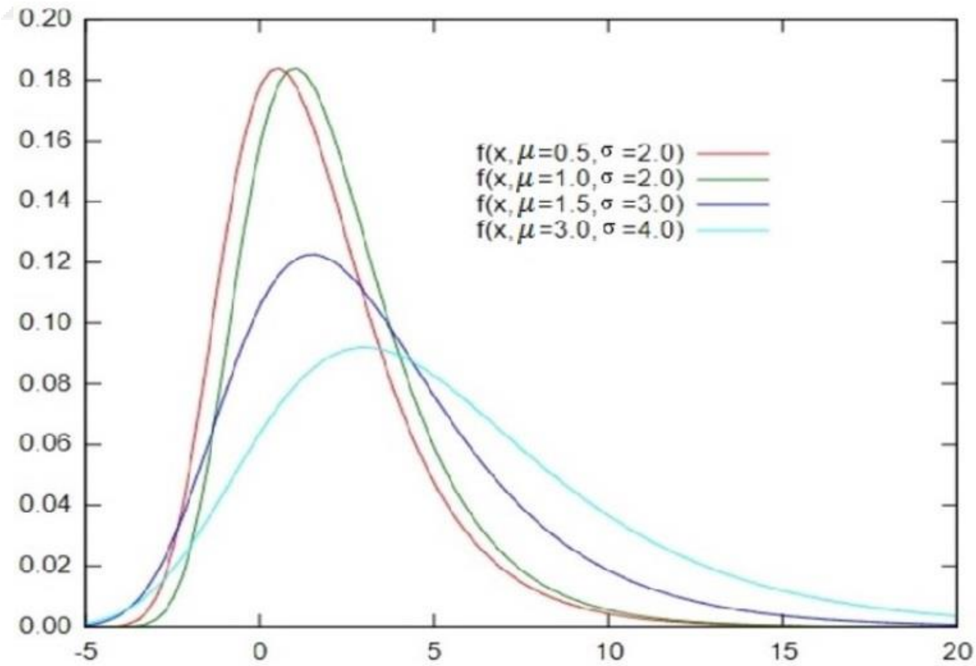

Figure 8: Performance the pdf s of Gumbel distribution.

In Figure 8, we can see the effective on probability density functions of Gumbel distribution when location and scale parameters take different values.

3.10 Burr type XII distribution

Burr type XII distribution is a member of a system of continuous distributions introduced by Irving w. Burr (1942).The importance of it is to model a heavy tailed data in simple way when compared with other distributions. Additionally, it has a greater flexibility, which enables it to be used by many applications. For instance, in deduction premium modeling see Burnecki, K., Härdle, W. and Weron, R. (2004). On the other hand, the three parameters Burr type XII distribution has been applied in time modeling for the survival of breast cancer patients in Gaza - Palestine for more information see Okasha, M. and Matter, M., (2015).

In the same context, The Burr Type XII distribution can be used to model the lifetime data see Feroze, N., and Aslam M., (2013). Recently, Burr family distribution have been used to characterize wind speed see Ouarda, T. B., Charron, C., and Chebana, F. (2016), Chiodo, E. and De Falco, P., (2016), Barcale, A., Carpinelli, G. and De Falco, P. (2017). The pdf as well as cdf written in the following way:

$$f(x) = \frac{ck}{\alpha}\ (\frac{x}{\alpha})^{c-1}\ [1 + \left(\frac{x}{\alpha}\right)^{c}]^{-k-1} \tag{3.36}$$

$$F(x) = 1 - \left[1 + \left(\frac{x}{\alpha}\right)^{c}\right]^{-k} \tag{3.37}$$

where c >0 and k >0 are the shape parameters and α is the scale parameter of the distribution. To find the maximum likelihood estimates (MLE) of parameters, which are unknown, ln L function is drafted as next:

$$\ln L = nlnc + nlnk - ncln\alpha + (c-1)\sum_{i=1}^{n} lnx_i - (k+1)\sum_{i=1}^{n}(1 + \frac{{x_i}^{c}}{\alpha^{c}}) \tag{3.38}$$

Next, derivatives are taken for ln L function for μ as well as λ . Afterwards, the equations will be equalized to zero. As a result, ML equations has been obtained in the following way:

$$\hat{k} = \frac{n}{k} - \sum_{i=1}^{n} (1 + (\frac{x_i}{\alpha})^c) \tag{3.39}$$

$$\hat{c} = \frac{n}{c} - nln\alpha + \sum_{i=1}^{n} lnx_i - (k+1)[\sum_{i=1}^{n} [\frac{(\frac{x_i}{\alpha})^c}{1 + (\frac{x_i}{\alpha})^c}] ln\frac{x_i}{\alpha}] \tag{3.40}$$

$$\hat{\alpha} = \frac{-nc}{\alpha c} - \frac{c(k+1)}{\alpha} \left[\sum_{i=1}^{n} \frac{\left(\frac{x_i}{\alpha}\right)^c}{1 + \left(\frac{x_i}{\alpha}\right)^c} \right] \tag{3.41}$$

Here, we solve Equations (3.39), (3.40) and (3.41) simultaneously by using iterative techniques.

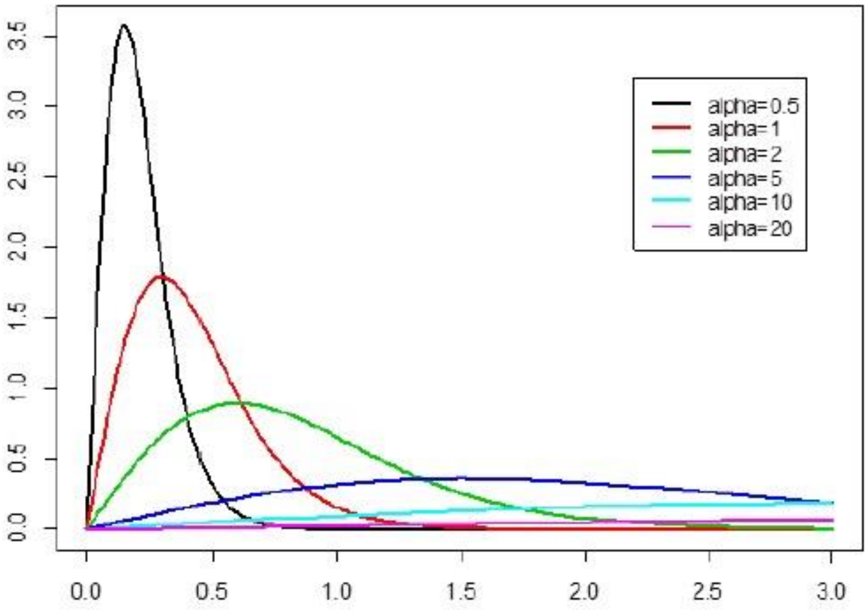

Figure 9: Performance the pdf s of Burr type XII distribution.

By checking figure 9, we can see the performance of distribution when scale parameter (α) getting higher and the shape parameters (c and k) kept as a constant (c=5, k=2). Following that, the distribution gets stretched out to the right and its height decreases, while maintaining its location. When the scale parameter went down and kept the shape parameters kept constant, it will lead that, the distribution is getting rise and pushes it to the left.

CHAPTER 4: MODEL EVALUATION

In following section, we will consider the selection criteria's which are used for the purpose of determining the probability density function which gives as the suitable model for the data of wind speed.

1. The first criteria is R square (coefficient of determination), It is the most common model from other models which is used to measure the amount of variance of observe data, which is explained by model. The equation of R square as bellows:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{F}_i - \bar{\bar{F}})^2}{\sum_{i=1}^{n}(\hat{F}_i - \bar{\bar{F}})^2 - \sum_{i=1}^{n}(\hat{F}_i - \frac{i}{n+1})^2}, \tag{4.1}$$

2. The second criteria is root mean square error, It is one of the widely selection forms used to examine the suitability of a pdf for the data. RMSE is used to measure the difference between the observe and predicted value. The equation of RMSE as follows:

$$RMSE = \left[\frac{1}{n}\sum_{n}^{i=1}(\hat{F}_i - \frac{i}{n+1})^2\right]^{1/2}, \tag{4.2}$$

3. The third criteria is Akaike information that is used to estimate the distance between unknown likelihood function of our data and the fitted likelihood function of our model. AIC equation as bellows:

$$AIC = -2lnL + 2p\,, \tag{4.3}$$

4. The fourth criteria is Bayesian information that is used to test which distribution has the suitable model for the data of wind speed. It worth to note that the BIC corrects the negative likelihood function by adding the number of estimated parameters multiplied by the logarithmic function of the sample size. On the other hand, AIC depends primarily on the number of estimated parameters as well as on the estimated negative likelihood function. The BIC equation as follows:

$$BIC = -2lnL + pln\,n, \tag{4.4}$$

5. Finally, Kolmogorov-Smirnov (KS) test has been developed by Kolmogorov in 1933 and Smirnov in 1939. It is used extensively in goodness of fit many problem such us suitability of fit. The Kolmogorov-Smirnov test used to calculate the biggest difference between the observing and forecasting distributions. The KS equation written as follows:

$$KS = \max_{1 \le i \le n} \left| \widehat{F}_i - \frac{i}{n+1} \right|, \tag{4.5}$$

where $\widehat{F}_i$ will be the estimated cdf, $\bar{\widehat{F}} = \frac{1}{n} \sum_{i=1}^{n} \bar{\widehat{F}}$, n will be the sample size, p will be the number of parameters which are estimated. It is important to be aware that for R^2 the higher value should be stated for best modeling. On the other hands, the lower value for RMSE, AIC, BIC, and KS should be stated for best fit. For further details, check Ouarda, T. B., Charron, C., and Chebana, F. (2016), and Dookie et al. (2018).

CHAPTER 5: CASE STUDY

Iraq is a suitable region for wind energy investment. In terms of location and climate, Iraq is located in the Arabian Peninsula, which is within the low-pressure area and under the influence of the Siberian airspace extension from the northern region through Turkey and from the northeast and eastern side through Iran in winter. In addition, Iraq is subject to the influence of the low air, which is semi-stable. This semi-stable airflow, which came from northwest India and central Asia towards the north and northwest part in Iraq during summer season. Therefore, these characteristics make Iraq an attractive area to winds. Additionally, the renewable energy that is based on the wind will decrease the reliant on fossil sources drastically.

The geographical layout in Iraq is interesting, because of the astronomical location, as well as with regard to land (land mass known as Eurasia). Thus, merits help the authorities' invest heavily in the winds energy.

On the other hand, In theory, Iraq is surrounded by five seas, represented by the Caspian Sea from the northeast, the Black Sea in the north, the Mediterranean Sea in the west, the Red Sea in the southwest, the Arabian Gulf and the Arabian Sea, where Iraq has a coast estimated at 58 km on the Arabian Gulf.

As previously noted, the wind energy in Iraq has a very good prospect Darwish, A. and Sayigh, A. M. (1988) this is because of the distinguished location of Iraq, its geographical location, its astronomical location, and its location in relation to land and water . As a result, the stations in Baghdad, Najaf, Hilla, Amara-Hai, Diwaniya, Nasiriya, Amara, and Basra from Iraq will be studied.

5.1 Dataset of wind speed

The average wind speed observation have been captured from Iraqi Meteorological Organization and Seismology – Ministry of Transportation, daily for the previous ten years starting from 2008 to 2018 at the heights of 10 m. The map is depicted in figure 10, which indicate the stations that are dispersed as well as close from each other's.

5.2 Geographical and Statistical Information

We can recognize the information, which concerns the geography of the reigns in Table 1, such us height from the surface of the sea, longitude, and latitude. Table 2, provides some important descriptive statistics which are calculated for the wind speed especially with regards mean, variances, maximum and number of observation, kurtosis, and skewness.

In chart 1, it reveals that Basra has attained the first position according the average speed of the wind. Regarding the average speed of the wind, Najaf station has the last position. On the other hand, the average of wind speed in stations Hilla, Diwaniya and Amara –Hai were less than the average of the wind speed in Nasiriya and Amara.

With respect to the variance chart 1 of wind speeds, the highest variance was observed in Basra and succeeded by Amara and Amara-Hai. At the same time, the lowest value of variance is observed in Najaf station. Consequently, the wind speed for this station is monolithic.

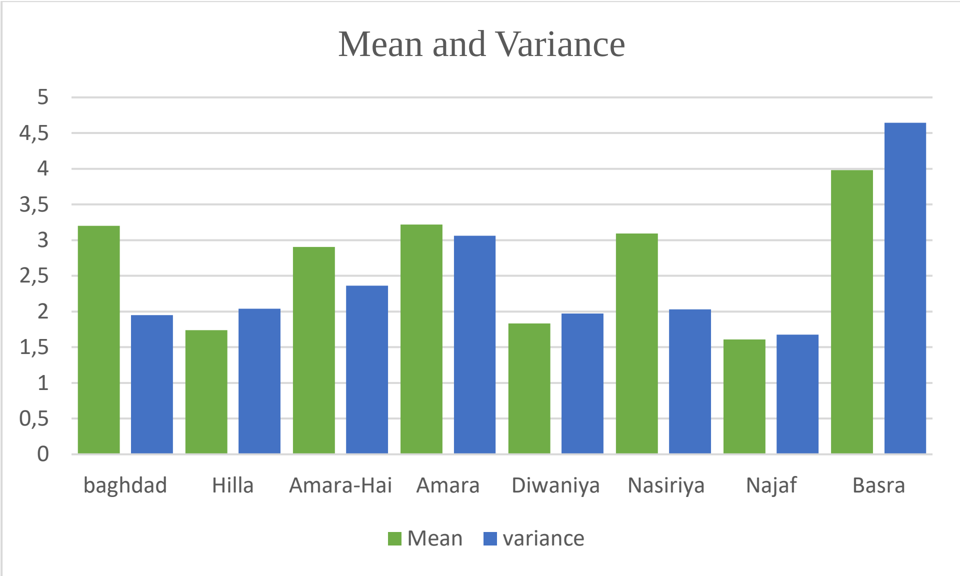

Chart 1: Mean and Variance for all stations.

Chart 2 has exhibited as anticipated a positive value of skewness. More spasificly, that Amara has attained the first position and Basra station has the last position according the skewness. Moreover, the values of skewness for Hilla and Diwaniya are greater than one while the skewness value for Baghdad, Amara-Hai, Nasiriya, and Najaf are less than one.

Regarding kurtosis chart 2, all the stations presented kurtosis with values less than three, which mean the distributions of the wind speeds are Platykurtic. On the other hand, wind speed data that has been recorded at Basra and Najaf has the smallest kurtosis values, while the speed of the wind data in Amara and Hilla has got the biggest value for the kurtosis respectively.

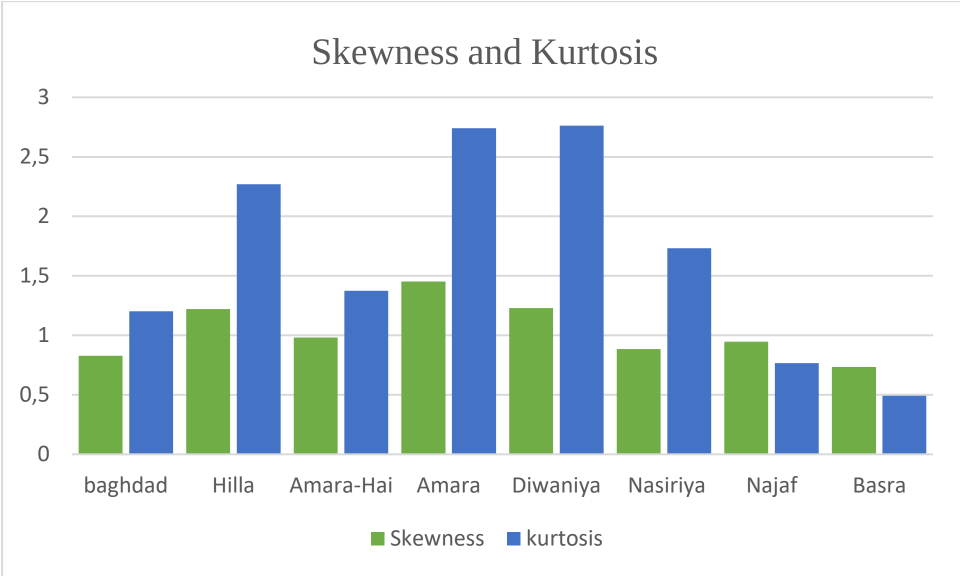

Chart 2: The Skewness and Kurtosis for all stations.

Regarding the maximum value of wind speed in chart 3, the highest result was observed in Basra. On the other hand, Najaf has the lowest regarding the maximum value of the wind speed.

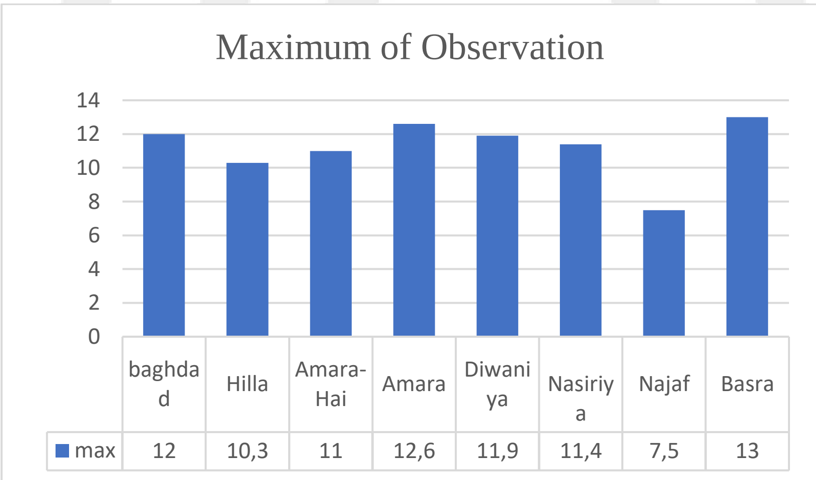

Chart 3: The Maximum of Observation for all stations.

Figure 10: Locations for stations under investigation (Source: Google map).

Table 1: The numerical geographical values for the stations.

Stations	Latitude	Longitude	height(m)
Baghdad	33 18	44 24	31.7
Najaf	31 57	44 19	53
Hilla	32 27	44 27	27
Amara-Hai	32 08	46 02	17
Diwaniya	31 57	44 57	20
Nasiriya	31 01	46 14	5
Amara	31 50	47 10	9.5
Basra	30 31	47 47	2

Table 2: Statistical values of the wind speed of our station under investigation.

Station	Mean	variance	Skewness	kurtosis	max	n	years
Baghdad	3.1992281	1.9493867	0.8269595	1.2026706	12	4016	2008-2018
Hilla	1.7402688	2.0398418	1.2200897	2.2708627	10.3	4018	2008-2018
Amara-Hai	2.9033375	2.363656	0.9805087	1.374023	11	4015	2008-2018
Amara	3.2186857	3.0638931	1.4515589	2.739842	12.6	3987	2008-2018
Diwaniya	1.8326109	1.9704962	1.2287695	2.7616458	11.9	4014	2008-2018
Nasiriya	3.0925206	2.0317795	0.8838016	1.7327211	11.4	4011	2008-2018
Najaf	1.6075162	1.6757812	0.9453684	0.7665373	7.5	4018	2008-2018
Basra	3.9815274	4.6462334	0.7338766	0.4919403	13	3627	2008-2018

CHAPTER 6: RESULT AND DISCUSSIONS

First, the maximum likelihood is acquired to estimate the unknown parameters of our set distributions for wind speed datasets, which are collected at Baghdad, Najaf, Hilla, Amara- Hai, Diwaniya, Nasiriya, Amara, and Basra stations in Iraq. These stations were taken for the purpose of comparison regarding the performance of wind speed. In our research, we take into account Weibull, Generalized Extreme value (GEV), Lognormal, Gamma, Rayleigh Burr type XII, Gumbel, and Inverse Gaussian, distributions to model the speed of the wind optimally.

Following that, Table 3 reveals the values, which are calculated of estimated parameters. On the other side, the results of the tests that have been found and used in our investigations such us, maximum likelihood function, R square, Akaike information criterion, root mean square error (RMSE), Bayesian information criterion, Kolmogorov-Smirnov test are presented in Table 4. In the same context, Figures 11-18 are depicted with probability density functions that they acquired by the model selection criteria to see the performance of distributions. Figures 19-26 will compare the attribute of Weibull distribution with the most promising one.

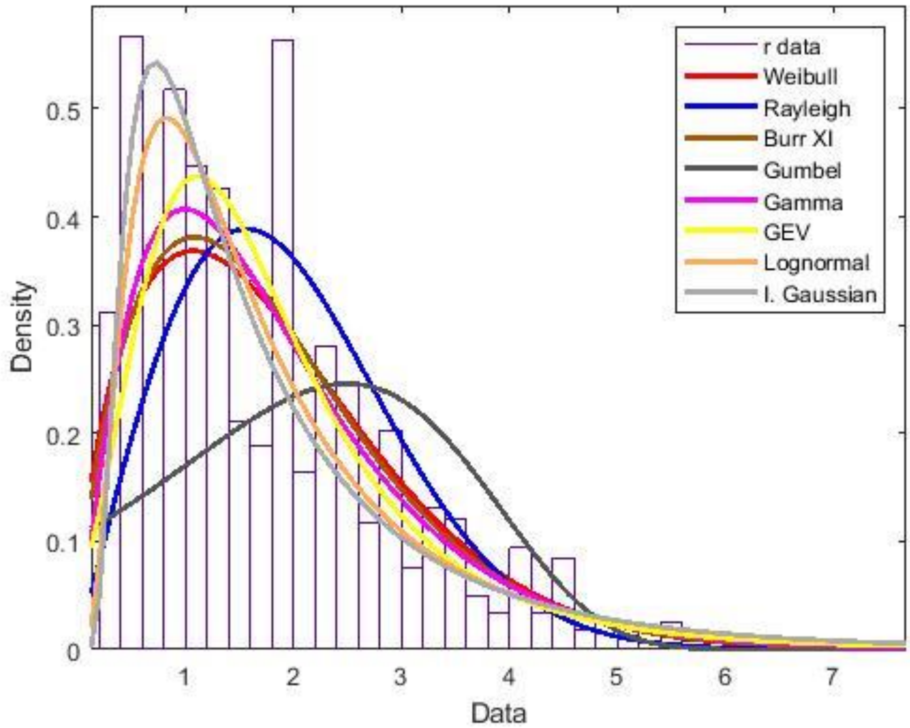

Figure 11: The histogram of Najaf station with suited densities.

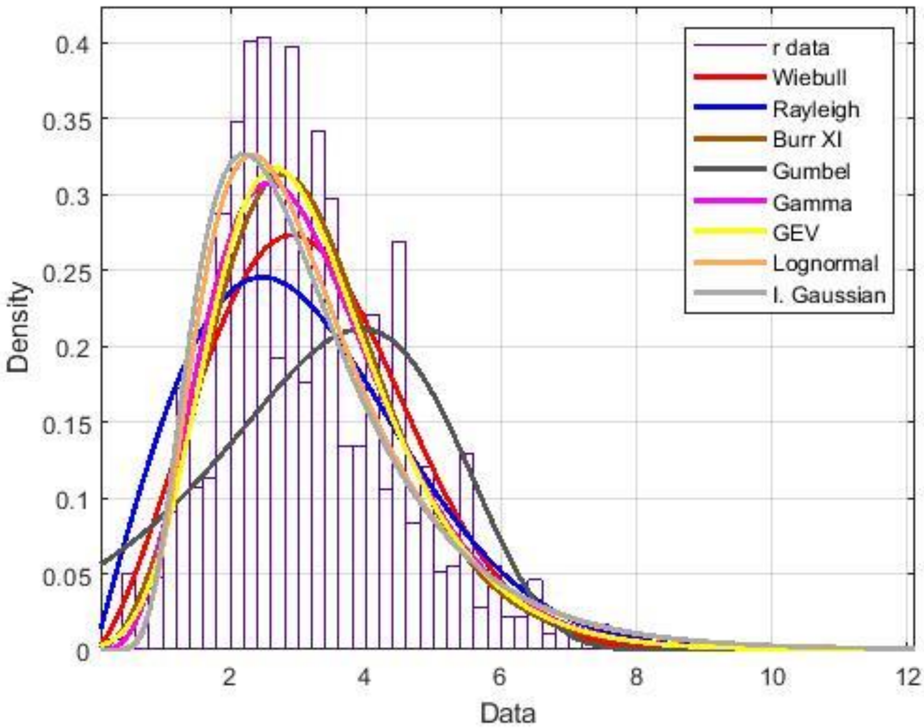

Figure 12: The histogram of Baghdad station with suited densities.

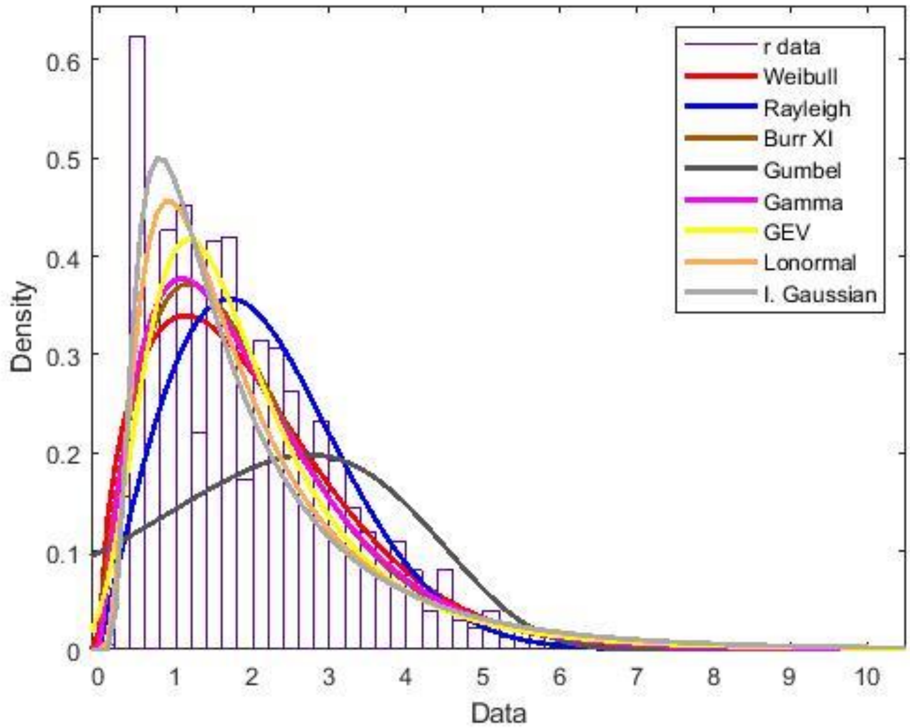

Figure 13: The histogram of Hilla station with suited densities.

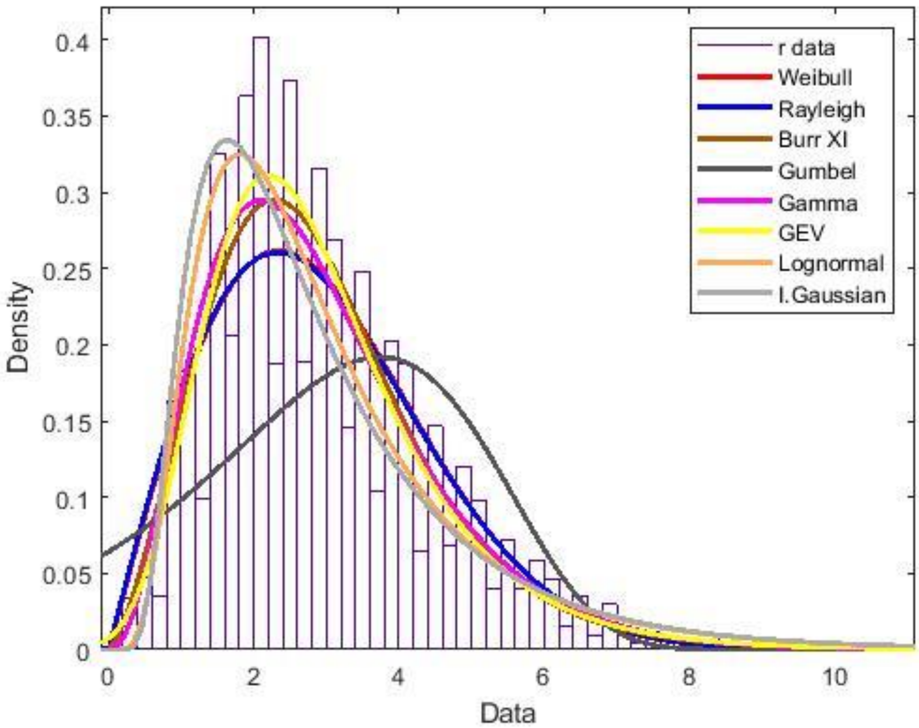

Figure 14: The histogram of Amara- Hai station with suited densities.

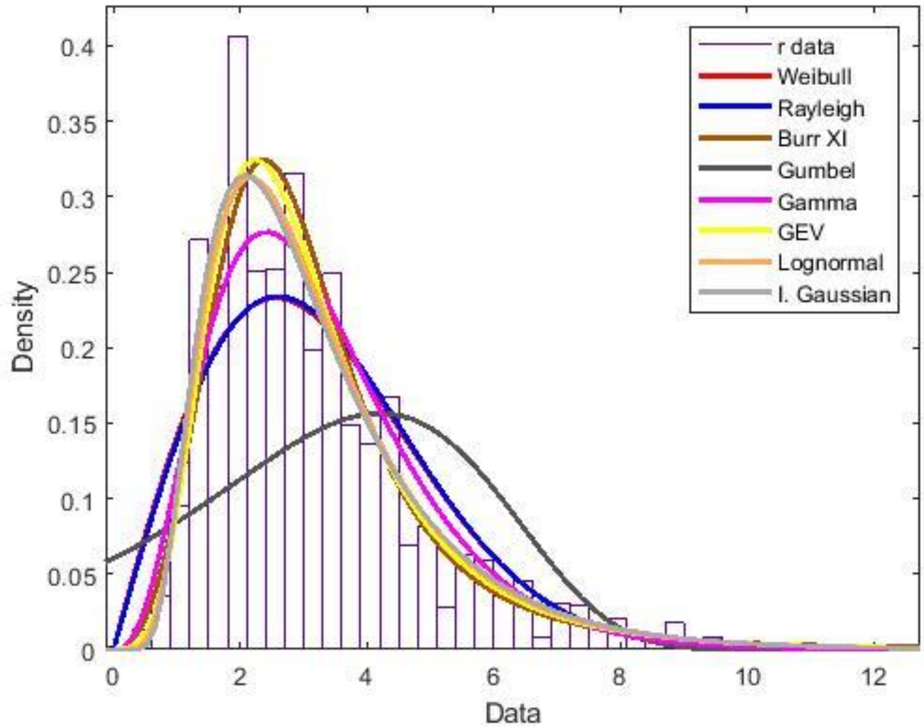

Figure 15: The histogram of Amara station with suited densities.

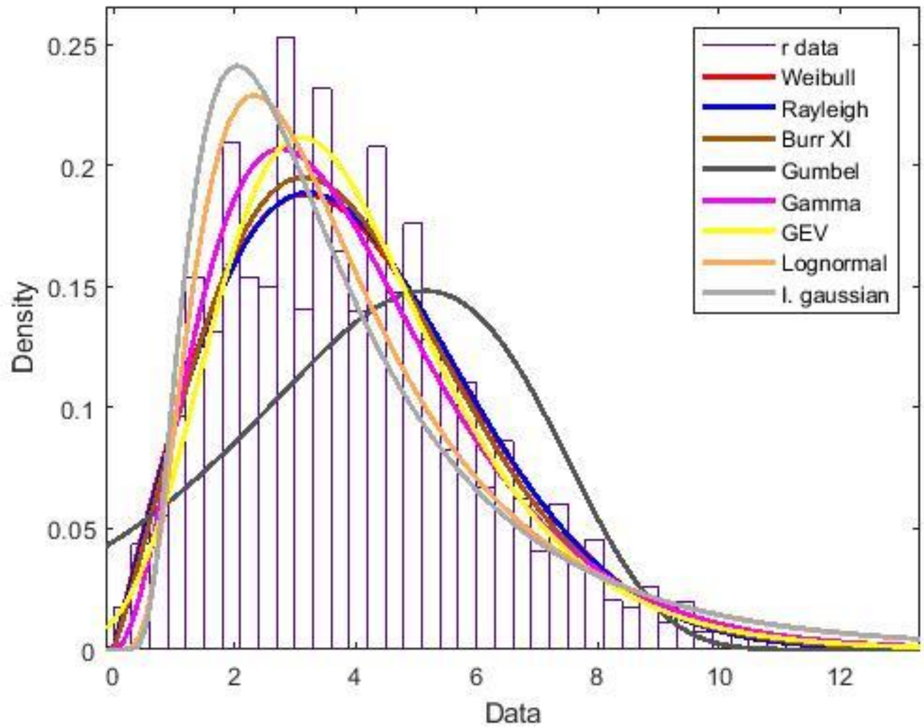

Figure 16: The histogram of Basra station with suited densities.

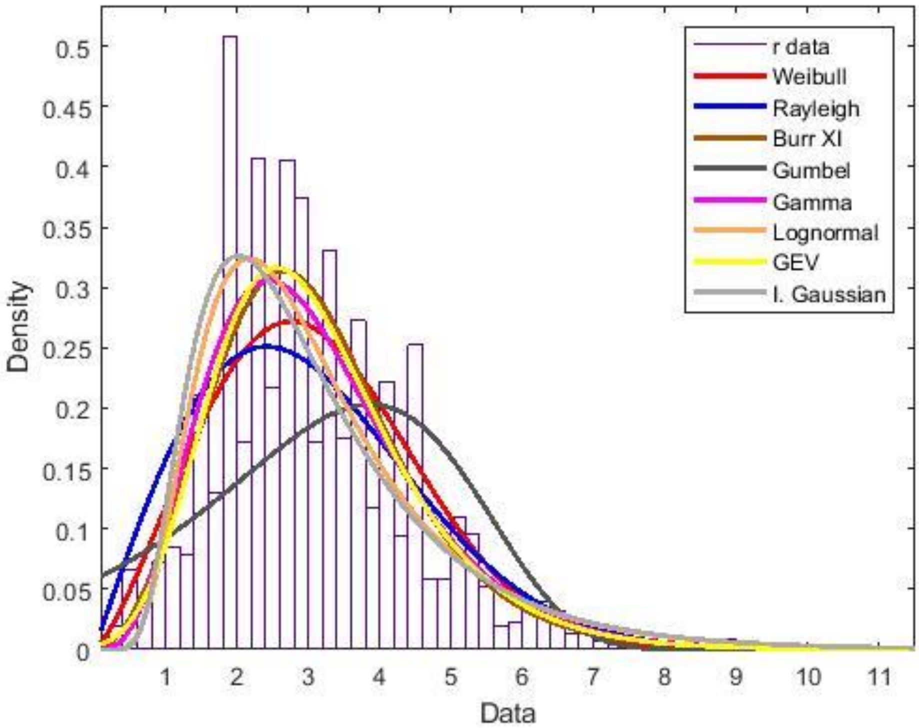

Figure 17: The histogram of Nasiriya station with suited densities.

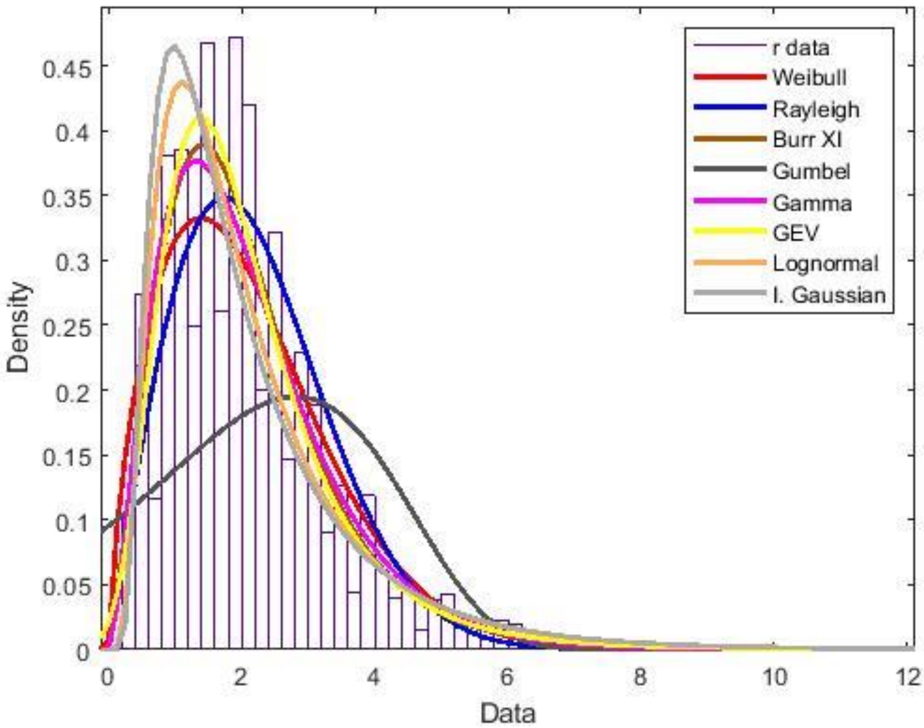

Figure 18: The histogram of Diwaniya station with suited densities.

Through the look at the Figures 11-18, we can notice that Weibull might not be the best fit in all stations and that each station has its suitable distribution.

In Najaf station, the results shows that, Lognormal distribution was performing more efficient with BIC, AIC and ln L than the remaining distributions. In addition to that, Weibull distribution has a good logical performance with R^2, while Gumbel distribution has reasonable performance with RMSE. On the other hand, the GEV distribution ranked firstly with KS test. The remaining distributions have unsatisfactory results in this station.

Following that, in Baghdad station, we found Lognormal distribution has an effective performance based on the result of RMSE and KS tests, while Burr distribution showed a strong performance according to R^2 and demonstrates superior execution as stated in lowest value of AIC, BIC and the highest value of ln L. On the other hand, the residual distributions have less performance for modeling wind speed in Baghdad station.

According Hilla station, it is easy to realize that Weibull has a good performance with the highest R^2. Furthermore, the lowest values of AIC, BIC and KS test was observed in Lognormal distribution. In addition, the results show that the Gumbel and Inverse Gaussian have a reasonable performance inconformity with RMSE and ln L respectively.

In Diwaniya station, the Lognormal obviously has an excellent performance to the lowest of, AIC, BIC, and the highest value of ln L, whereas, Weibull performs was better regarding of R^2. In the same station, the results appear that Gumbel distribution has the best attribute based on RMSE value, while GEV distribution get the first rank according to KS test.

In Amara- Hai station, lognormal demonstrates reasonably great modeling performance with highest ln L and lowest of AIC, BIC, while the Gumbel introduced good performance with lowest RMSE, whereas Rayleigh distribution shows better performance with KS test. Moreover, GEV distribution takes top rate based on the highest value of R^2.

Finally, Amara, Basra and Nasiriya stations, they displayed clearly an outstanding performance than the other distributions regarding Lognormal in accordance with lowest AIC, BIC and highest ln L. Interestingly, the outcomes of Amara station displayed that the Gumbel shows better performance with RMSE, besides that, Burr has revealed a good performance according to the highest R^2 and the best result of KS test went to Rayleigh distribution.

On the other hand, Basra station constituted that Gumbel has acquired a second rank of performance based on RMSE and R^2, while GEV has preferable performance due to the result of KS test. On the contrary, Nasiriya station portrayed that Rayleigh and Gumbel have a significant performances because of KS and RMSE values. Burr also reveals best performance with R^2.

It can be argued that the wind speed characteristics that we have observed through our experiment might relatively be different from each other regardless the locations adjacency. Based on the foregoing interpretation of the results in Table 4, Weibull Distribution is commonly used to explain most models of wind speed, though its performance was not distinguished in most stations under investigation. As noted, The Lognormal ranked first in its performance for wind speed modeling, -especially in Diwaniya, Najaf , Basra, Amara, Amara- Hai, Nasiriya, and Hilla- where it achieved the best results in most tests as the best performance.

Furthermore, the performance of Burr type XII Distribution dominated on all distributions according to the values of (R^2, ln L, AIC, and BIC) in Baghdad station.

Weibull, burr type XII and Generalized Weibull distributions, have achieved significant performance in wind speed modeling based on the results of R^2. On the other hand, Rayleigh, GEV and Lognormal, their performance in wind speed modeling was respectable according to KS test values.

In the end, the distributions Gamma, and Inverse Gaussian, did not perform well enough for modeling wind speed in the environment that we have tested for our case study.

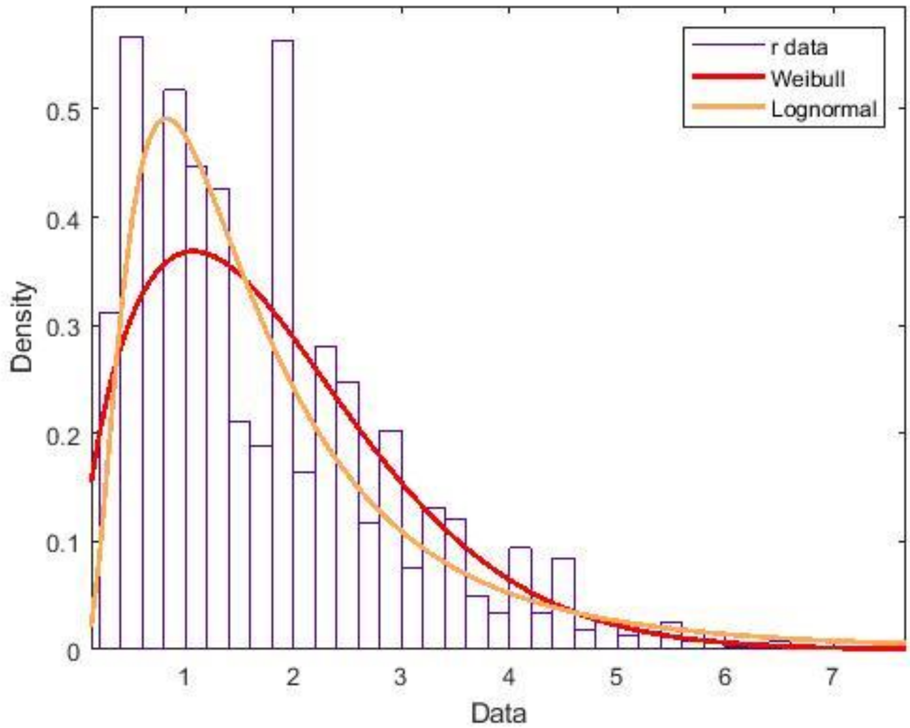

Figure 19: The performance of Weibull with lognormal distributions in Najaf station.

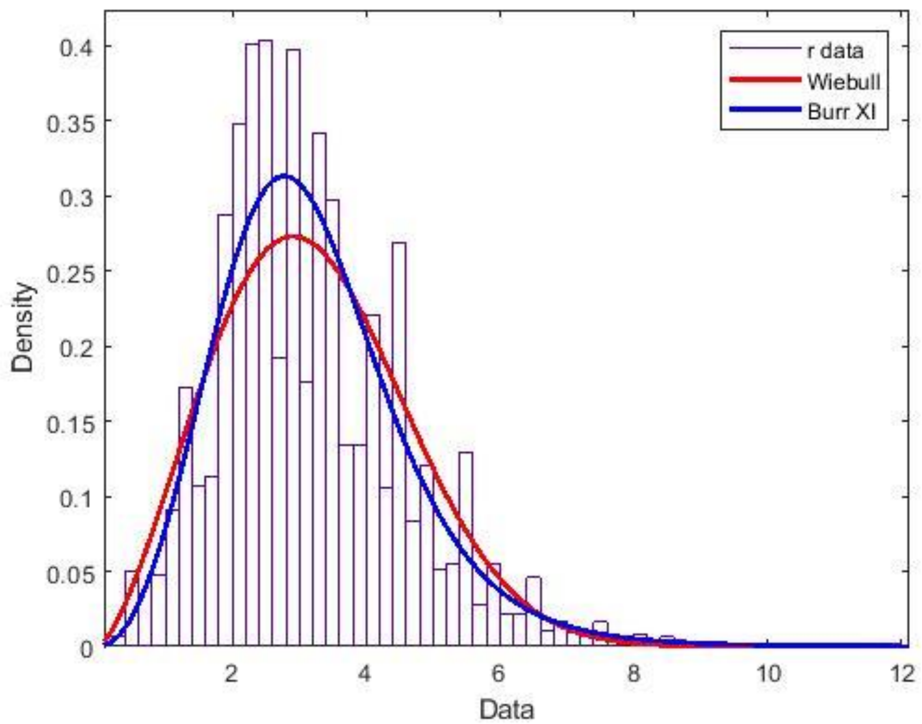

Figure 20: The performance of Weibull with Burr type XII distributions in Baghdad station.

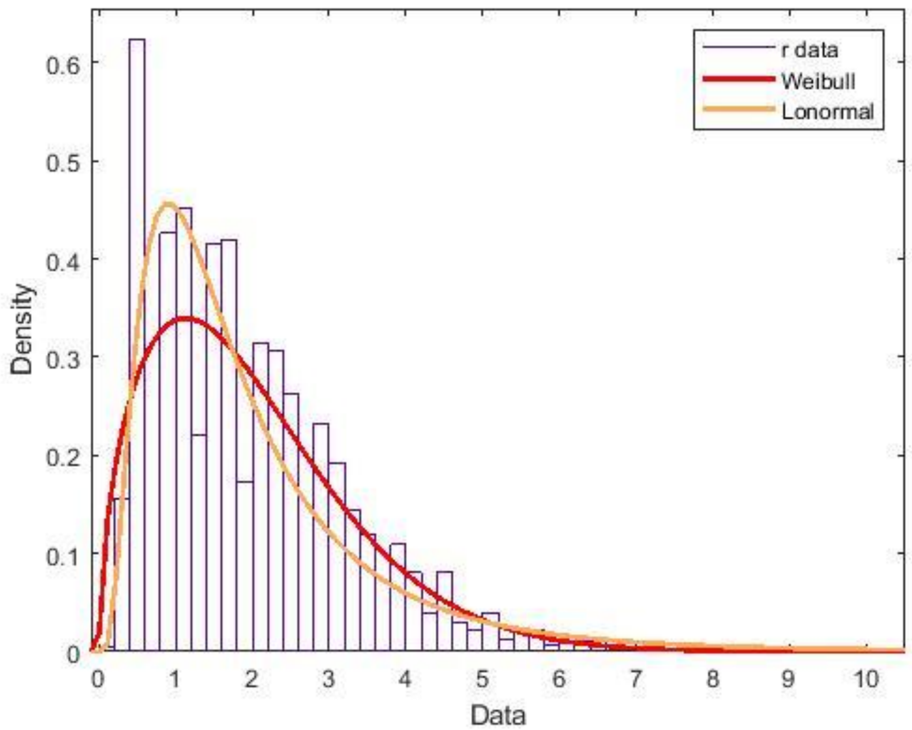

Figure 21: The performance of Weibull with lognormal distributions in Hilla station.

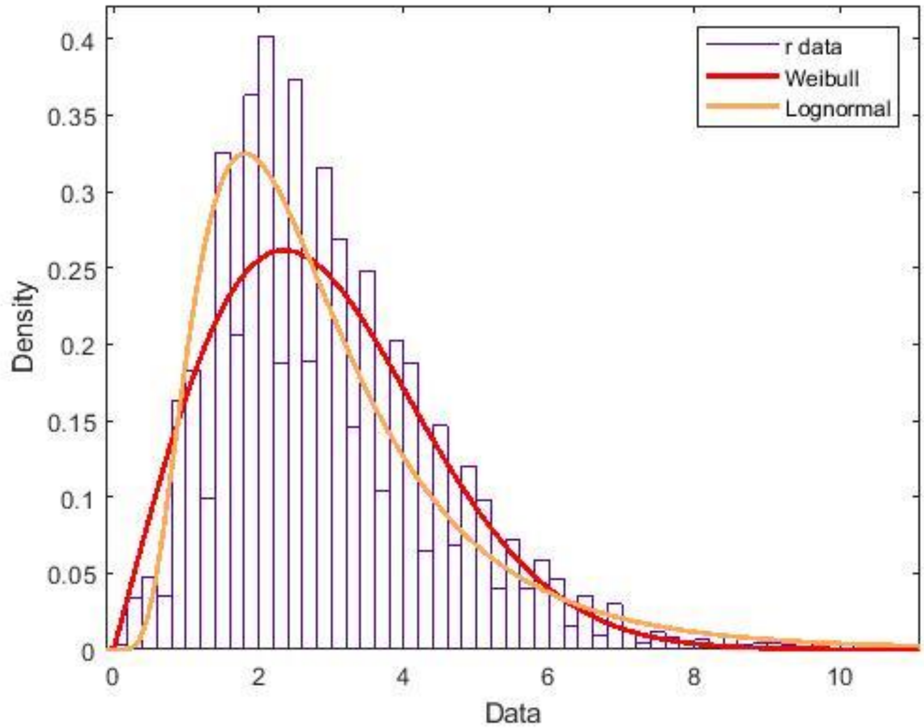

Figure 22: The performance of Weibull with lognormal distributions in Amara- Hai station.

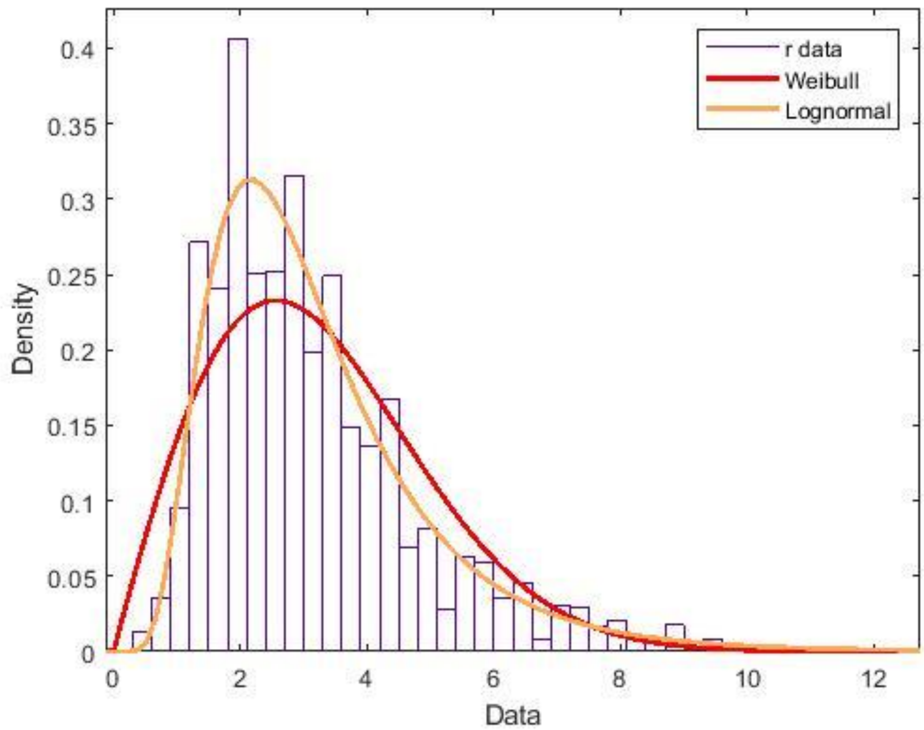

Figure 23: The performance of Weibull with Lognormal distributions in Amara station.

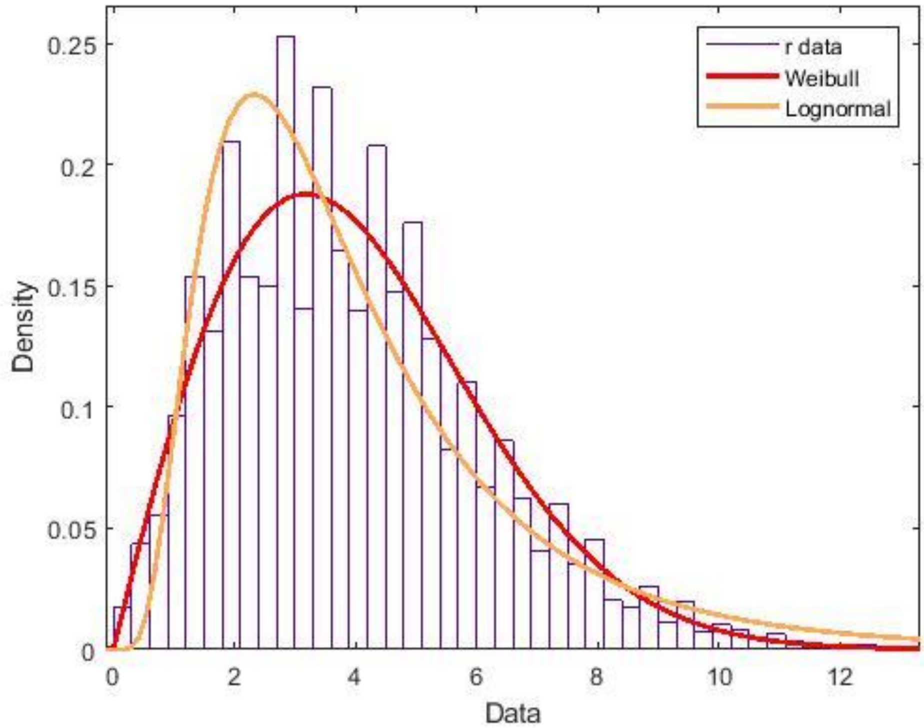

Figure 24: The performance of Weibull with Lognormals distribution in Basra station.

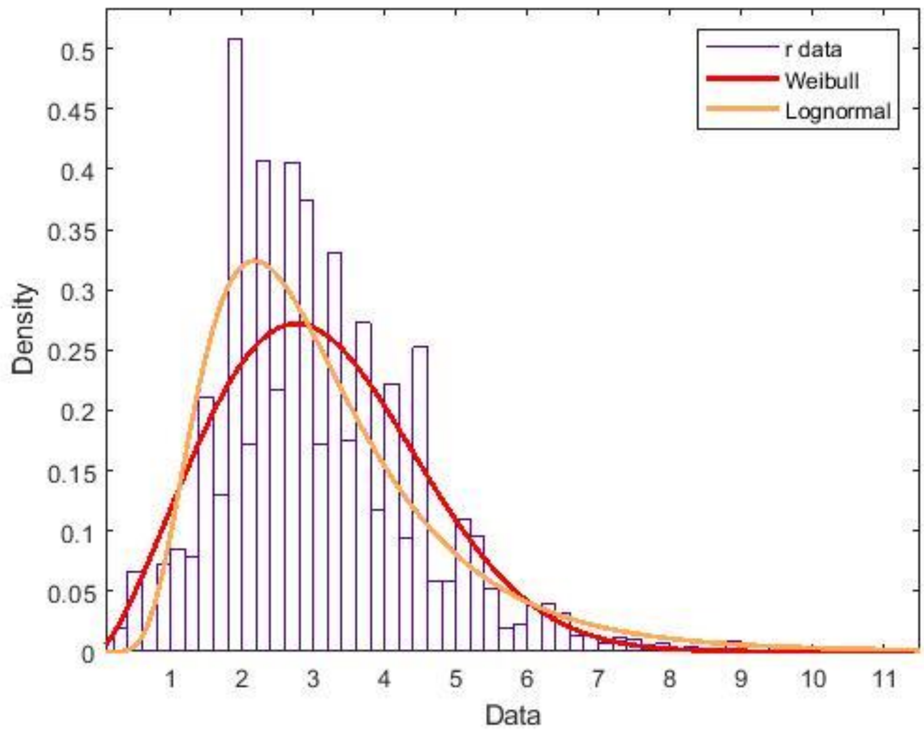

Figure 25: The performance of Weibull with Lognormal distributions in Nasiriya station.

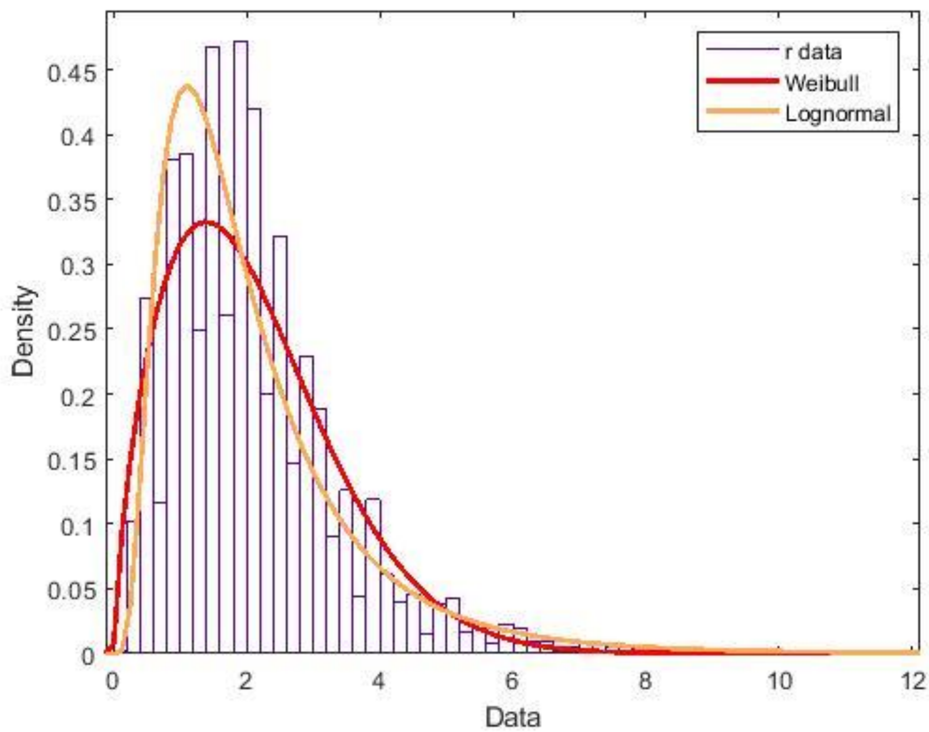

Figure 26: The performance of Weibull with Lognormal distributions in Diwaniya station.

Table 3: The maximum likelihood estimates of parameters of wind speed distributions.

Distributions	parameter	Stations							
		Najaf	**Baghdad**	**Hilla**	**Amara-hai**	**Diwaniya**	**Nasiriya**	**Amara**	**Basra**
Weibull distribution.	shape(c)	1.568146	2.431828	1.54525	2.0188	1.6956	2.3271	1.9770	1.9774
	scale (σ)	2.050	3.614	2.2135	3.2999	2.3463	3.5145	3.6568	4.5342
Rayleigh distribution.	scale (σ)	1.5594	2.4697	1.6990	2.3285	1.7410	2.4159	2.5933	3.2142
Burr type XII dist.	alpha	8.5284	4.5944	4.5404	4.8603	3.1839	4.5358	2.9112	13.2866
	shape(c)	1.6505	3.0386	1.7771	2.4291	2.1562	2.9256	3.3185	2.0968
	shape (k)	11.2523	2.7717	4.3036	3.2697	2.6275	2.8092	1.0704	10.2642
Gumbel dist.	location (м)	2.4976	3.9907	2.7313	3.7415	2.8093	3.8671	4.1990	5.1487
	scale (σ)	1.4965	1.7393	1.8575	1.9174	1.8862	1.8124	2.3509	2.4835
Gamma distribution.	shape(c)	2.1841	5.1250	2.2009	3.5676	2.6803	4.6897	3.9516	3.2133
	scale (σ)	0.8404	0.6250	0.9008	0.8177	0.7779	0.6639	0.8162	1.2491
Generalized extreme value dist.	shape(k)	0.1408	0.0541	0.1692	0.0128	0.1007	0.0480	0.1424	0.0309
	scale (σ)	0.8483	1.1598	0.8932	1.1825	0.9000	1.1638	1.1437	1.7397
	location (м)	1.2200	2.5928	1.3040	2.2166	1.4684	2.4980	2.3846	3.0533
Lognormal dist.	mean (м)	0.3612	1.0634	0.4403	0.9246	0.5368	1.0254	1.0391	1.2261
	S. D.(σ)	0.7476	0.9713	0.7382	0.5748	0.6636	0.4999	0.5150	0.6191
inverse Gaussian dist.	scale (м)	1.8355	3.2032	1.9825	2.9171	2.0851	3.1135	3.2252	4.0136
	shape (σ)	2.5334	12.4880	2.7969	7.1300	3.7297	10.4180	10.6050	8.2601

Table 4: List of criteria's (R^2, lnL, AIC, BIC, RMSE, and KS) values for the wind speed distributions.

Station	Criteria	Weibull	Rayleigh	Burr XII	Gumbel	Gamma	GEV	Lognormal	I. Gaussian
Najaf	RMSE	0.4792	0.4611	0.4384	0.3434	0.4397	0.4432	0.4372	0.4362
	R^2	0.3354	0.3233	0.3108	0.3141	0.3126	0.3166	0.3119	0.3104
	lnL	-5176.5	-22665	-5172.7	-6942.9	-5155	-25120	-3481.3	-5264.4
	AIC	10357	45332	10351	13890	10314	50247	6966	10533
	BIC	10369	45338	10370	13902	10326	50265	6978.9	10545
	KS	0.9978	0.9847	0.9713	1.9678	0.9680	0.9607	0.9804	0.9853
Baghdad	RMSE	0.4745	0.3716	0.3970	0.3785	0.3938	0.3926	0.3258	0.3871
	R^2	0.3011	0.3117	0.3467	0.2649	0.3433	0.3413	0.2137	0.3345
	lnL	-6874.1	-152210	-6800.1	-8619.2	-6809.7	-75002	-7834.5	-7026
	AIC	13752	304430	13606	17242	13623	150010	15673	14056
	BIC	13765	304430	13625	17255	13636	150030	15686	14069
	KS	0.9985	0.9913	0.9901	1.9985	0.9910	0.9927	0.9066	0.9968
Hilla	RMSE	0.4629	0.4288	0.4098	0.3571	0.4093	0.4136	0.4088	0.4077
	R^2	0.3574	0.3489	0.3400	0.2557	0.3395	0.3441	0.339	0.3366
	lnL	-5476.1	-31575	-5447.6	-7221.9	-5430.1	-21288	-3723.5	-3520.3
	AIC	10956	63153	10901	14444	10864	42583	7451	11045
	BIC	10968	63159	10920	14460	10877	42601	7463.4	11057
	KS	0.9943	0.9943	0.9917	1.9610	0.9931	0.9878	0.9856	0.9901
Amara- Hai	RMSE	0.4781	0.3920	0.3993	0.3720	0.3969	0.4010	0.3921	0.3881
	R^2	0.3115	0.3351	0.3449	0.2705	0.3425	0.3472	0.3383	0.3326
	lnL	-7058.6	-120530	-7008.3	-8780.4	-7006.7	-312990	-5161.5	-7283.6
	AIC	14121	241070	14023	17565	14017	625880	10327	14571
	BIC	11434	241070	14023	17577	14030	625900	1033.9	14584
	KS	0.9985	0.9787	0.9832	1.9637	0.9853	0.9812	0.9869	0.9870

Table 4 (Cont'd): List of criteria's (R^2, lnL, AIC, BIC, RMSE, and KS) values for the wind speed distributions.

Station	Criteria	Weibull	Rayleigh	Burr XII	Gumbel	Gamma	GEV	Lognormal	I. Gaussian
Diwaniya	RMSE	0.4199	0.3930	0.3833	0.3575	0.3805	0.3841	0.3761	0.3731
	R^2	0.3763	0.3557	0.3614	0.2333	0.3574	0.3637	0.3592	0.3565
	lnL	-5449.6	-34434	-5372	-70970	-5378.1	-35187	-3689.1	-5524.8
	AIC	10903	68869	10750	14198	10760	70374	7382.2	11054
	BIC	10916	68875	10769	14210	10773	70398	7394.6	11066
	KS	0.9991	0.9912	0.9894	1.9912	0.9909	0.9893	0.9897	0.9918
Nasiriya	RMSE	0.4806	0.3948	0.4194	0.3464	0.4140	0.4141	0.4069	0.4020
	R^2	0.2944	0.2914	0.3187	0.2870	0.3149	0.3143	0.3112	0.3067
	lnL	-6850.9	-138670	-6767.4	-8440	-6800.8	-83844	-4983.8	-7112.2
	AIC	13706	277330	13541	16884	13606	167690	9971.7	14228
	BIC	13718	277430	13560	16897	13618	167710	9984	14241
	KS	0.9985	0.9894	0.9878	1.9612	0.9886	0.9845	0.9876	0.9871
Amara	RMSE	0.5207	0.4104	0.4293	0.3455	0.4230	0.4283	0.4255	0.4236
	R^2	0.2315	0.3004	0.3194	0.2670	0.3135	0.3181	0.3152	0.3129
	lnL	-7427	-183410	-7176.1	-9130.2	-7214.1	-29296	-5150.9	-7150.2
	AIC	14858	366820	14358	18264	14432	58599	10306	14304
	BIC	14871	366820	14377	18277	14445	58618	10318	14317
	KS	0.9985	0.9776	0.9929	1.9633	0.9920	0.9953	0.9969	0.9977
Basra	RMSE	0.5791	0.4756	0.4771	0.3024	0.4763	0.4789	0.4697	0.4656
	R^2	0.2111	0.2675	0.2689	0.3878	0.2686	0.2704	0.2637	0.2591
	lnL	-7588.1	-388020	-7582.9	-9298.3	-7602.3	-118580	-5992.6	-7924.1
	AIC	15180	776040	15172	18601	15209	237170	11989	15852
	BIC	15193	776050	15190	18613	15221	237190	12002	15865
	KS	0.9911	0.9911	0.9911	1.9284	0.9936	0.9884	0.9947	0.9947

6.1 Conclusion

Calculating the susceptibility and determining the characteristics of wind speed largely depends on the appropriate statistical distributions. In our study, 8 wind distributions were used as follows: Weibull, Rayleigh, Burr type XII, Gumbel, Gamma, Inverse Gaussian, Generalized Extreme value, and Lognormal. These distributions were used for their high ability to demonstrate wind speed modeling. We were clearly able to see that Weibull distribution, which is widely defined and widely used, remains unable to model all wind speed regimes. For these reasons, we have relied on the use of alternative distributions.

As it is mentioned in the previous section, from Table 4 it is easy to note that Lognormal distribution was the best model for wind speed performance, while Burr type XII distribution ranked second in its ability to model wind speed. Likewise, Weibull, Gumbel, Rayleigh and GEV distributions had a good and reasonable performance at most stations. The performance of Gamma, and Inverse Gaussian were not sufficient in explaining the wind speed regimes. Consequently, we reached a conclusion that each region has its own characteristics and some distributions, which capable to characterize the performance of wind speed, in each region appropriately.

REFERENCES

Acker, T., Williams, S., Duque, E., Brummels, G. and Buechler, J., (2007) *Wind resource assessment in the state of Arizona: Inventory, capacity factor, and cost*, Renewable Energy, vol. 32(9), pp.1453-1466.

Ahmed Shata A.S., and Hanitsch R. (2006) *Evaluation of wind energy potential and electricity generation on the coast of Mediterranean Sea in Egypt*, Renew Energy, vol. 31(8), pp.1183–202.

Ahmed, A. S. (2010) *Wind energy as a potential generation source are Ras Benas, Egypt*, Renew Sustain Energy, vol. 4:2167–73. doi:10.1016/j.rser.2010.03.006.

Akdağ, S. and Güler, Ö. (2018) *Alternative Moment Method for wind energy potential and turbine energy output estimation*. Renewable Energy, vol. 120, pp.69-77.

Akgul, F., Arslan, T. and Senoglu, B. (2016*) An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution*, Energy convers Manage, vol. 114, pp.234-240.

Akpinar EK, and Akpinar S. (2005) *A statistical analysis of wind speed data used in installation of wind energy conversion systems*, Energy Convers Manage, vol. 46(4), pp.515–32.

Alavi, O., Mohammadi, K. and Mostafaeipour, A. (2016) *Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran*, Energy Conversion and Management, vol. 119, pp.101-108.

Alavi, O., Sedaghat, A. and Mostafaeipour, A. (2016) *Sensitivity analysis of different wind speed distribution models with actual and truncated wind data: A case study for Kerman, Iran*, Energy Conversion and Management, vol. 120, pp.51-61.

Aldrich, J. (1997) *R.A. Fisher and the making of maximum likelihood 1912-1922,* Statistical Science, vol. 12(3), pp.162-176.

Al-Fawzan, M., (2000) *Methods for Estimating the Parameters of the Weibull Distribution*, King Abdulaziz City for Science and Technology [Online]. Available at https://www.researchgate.net/publication/240935449 (Accessed10 June 2020).

Al-Hasan, M. and Nigmatullin, R., (2003) *Identification of the generalized Weibull distribution in wind speed data by the Eigen-coordinates method,* Renewable Energy, vol. 28(1), pp.93-110.

Amaya-Martínez, P., Saavedra-Montes, A. and Arango-Zuluaga, E. (2014*) A statistical analysis of wind speed distribution models in the Aburrá Valley, Colombia*, CT&F - Ciencia, Tecnología y Futuro, vol. 5(5), pp.121-136.

Arslan, T., Acitas, S. and Senoglu, B. (2017) *Generalized Lindley and Power Lindley distributions for modeling the wind speed data*, Energy Conversion and Management, vol. 152, pp.300-311.

Bagiorgas, H., Mihalakakou, G., Rehman, S. and Al-Hadhrami, L. (2015) *Wind power potential assessment for three buoys data collection stations in the Ionian Sea using Weibull distribution function*, International Journal of Green Energy, vol. 13(7), pp.703-714.

Belikov, A., (2017) *The number of key carcinogenic events can be predicted from cancer incidence*. Scientific Reports, vol. 7(1).

Bidaoui, H., Abbassi, I., Bouardi, A. and Darcherif, A., (2019) *Wind Speed Data Analysis Using Weibull and Rayleigh Distribution Functions, Case Study: Five Cities Northern Morocco*, Procedia Manufacturing, vol. 32, pp.786-793.

Bilir, L., Devrim, Y. and Imir, M. (2015) *An investigation on wind energy potential and small-scale wind turbine performance at Incek region – Ankara, Turkey*, Energy Convers Manage, vol. 103, pp.910–923.

Black, F. and Scholes, M. (1973) *The Pricing of Options and Corporate Liabilities*. Journal of Political Economy, vol. 81 (3): 637. doi:10.1086/260062.

Bracale, A., Carpinelli, G. and De Falco, P., (2017) *A new finite mixture distribution and its expectation-maximization procedure for extreme wind speed characterization*, Renewable Energy, vol. 113, pp.1366-1377.

Brano, V., Giullaa, G., Orioli, A. and Culotta, S. (2011) *Quality of wind speed fitting distributions for the urban area of Palermo, Italy*, Renew Energy, vol. 1, p.14.

Burnecki, K., Härdle, W. and Weron, R. (2004) *An introduction to simulation of risk processes*, Encyclopedia of Actuarial Science [Online]. Available at https://www.researchgate.net/publication/240935449, pp.1–7, (Accessed10 January 2019).

Burr, I. (1942) *Cumulative frequency functions*, Annals of Mathematical Statistics, vol.13 (2):215–232. doi:10.1214/aoms/1177731607, JSTOR 2235756.

Carroll, K., (2003) *On the use and utility of the Weibull model in the analysis of survival data*, Controlled Clinical Trials, vol. 24(6), pp.682-701.

Carta, J., Ramírez, P. and Velázquez, S., (2009) *A review of wind speed probability distributions used in wind energy analysis*, Renewable and Sustainable Energy Reviews, vol. 13(5), pp.933-955.

Chang, T. (2011) *Estimation of wind energy potential using different probability density functions*, Applied Energy, vol. 88:1848–1856, pp.10.

Cheng, E. and Yeung, C. (2002) *Generalized extreme gust wind speeds distributions*, Journal of Wind Engineering and Industrial Aerodynamics, vol. 90(12-15), pp. 1657-1669.

Chiodo, E. and De Falco, P., (2016) *Inverse Burr distribution for extreme wind speed prediction: Genesis, identification and estimation*, Electric Power Systems Research, vol. 141, pp.549-561.

Costa Rocha, P., de Sousa, R., de Andrade, C. and da Silva, M., (2012) *Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil*, Applied Energy, vol. 89(1), pp.395-400.

Dahm, H., Niemeyer, J. and Schröder, D., (2002) *Application of the Weibull distribution to describe the vertical distribution of cesium-137 on a slope under permanent pasture in Luxembourg*, Journal of Environmental Radioactivity, vol. 63(3), pp.207-219.

Darwish, A. and Sayigh, A., (1988) *Wind energy potential in Iraq*, Solar & Wind Technology, vol. 5(3), pp.215-222.

Debanshee, D. and Datta, A. M. (2013*) Comparison of Weibull Distribution and Exponentiated Weibull Distribution Based Estimation of Mean and Variance of Wind Data*, International Journal of Energy, vol. 4(4), pp. 12.doi:10.1016/j.enconman.2016.04.078.

Dookie, I., Rocke, S., Singh, A. and Ramlal, C., (2018), *Evaluating wind speed probability distribution models with a novel goodness of fit metric: a Trinidad and Tobago case study*, International Journal of Energy and Environmental Engineering, vol. 9(3), pp.323-339.

Edgeworth, F., (1908) *On the Probable Errors of Frequency-Constants (Contd.)*, Journal of the Royal Statistical Society, vol. 71(4), p.651.

Feroze, N., and Aslam M., (2013) *Maximum Likelihood Estimation of Burr Type V Distribution under Left Censored Samples*, Wseas Transactions on Mathematics, vol. 12(6), pp. 13.

Fleming, R., (2001) *The Weibull model and an ecological application: describing the dynamics of foliage biomass on Scots pine*, Ecological Modelling, vol. 138(1-3), pp.309-319.

Gani, A., Mohammadi, K., Shamshirband, S., Altameem, T. A., Petkovic, D. and Ch, S. (2016) *A combined method to estimate wind speed distribution based on integrating the support vector machine with firefly algorithm*, Environmental Progress & Sustainable Energy, vol. 35 (3), pp. 867–75.

Ginos, B., (2009) *Parameter Estimation for the Lognormal Distribution*, Unpublished Doctoral Thesis, Department of Statistics Brigham Young University.

Guégan, D., and Hassani, B.K. (2014) *A mathematical resurgence of risk management: an extreme modeling of expert opinions*, Frontiers in Finance and Economics, vol. 11 (1): 25–45.

Gugliani, G., Sarkar, A., Mandal, S. and Agrawal, V. (2017) *Location wise comparison of mixture distributions for assessment of wind power potential: A parametric study*, International Journal of Green Energy, vol. 14(9), pp.737-753.

Gumbel E.J. (1941) *The return period of flood flows*, The Annals of Mathematical Statistics, vol. 12, 163–190.

Hassoon A., (2013) *Assessment potential wind energy in the north area of Iraq*, International Journal of Energy and Environment, vol. 4(5), pp.807-814.

Hepbasli, A. and Ozgener, O. (2004) *A review on the development of wind energy in Turkey*, Renewable and Sustainable Energy Reviews, vol. 8(3), pp.257-276.

Hogg, R., McKean, J. and Craig, A. (1978) *Introduction to Mathematical Statistics*. 4th edition. Edited by Robert V Hogg, Joseph W. McKean, and Allen T. Craig .New York: Macmillan.

Huillet, T. and Raynaud, H., (1999) *Rare events in a log-Weibull scenario - Application to earthquake magnitude data*, The European Physical Journal B, vol. 12(3), pp.457-469.

Ivana, P., Zuzana, S. and Mária, M. (2017) *Application of four probability distributions for wind speed modeling*, Procedia Engineering, vol. 192, pp. 713 – 718.

Jowder, F. (2009) *Wind power analysis and site matching of wind turbine generators in Kingdom of Bahrain*, Applied Energy, vol. 86(4), pp.538-545.

Jung, C. and Schindler, D. (2017) *Global comparison of the goodness-of-fit of wind speed distributions*, Energy Conversion and Management, vol. 133, pp.216-234.

Kang, D., Ko, K. and Huh, J., (2015) *Determination of extreme wind values using the Gumbel distribution*. Energy, vol. 86, pp.51-58.

Kantar, Y. and Usta, I. (2015) *Analysis of the upper-truncated Weibull distribution for wind speed*, Energy Conversion and Management, vol. 96, pp.81-88.

Khahro, S., Tabbassum, K., Soomro, A., Dong, L. and Liao, X. (2014) *Evaluation of wind power production prospective and Weibull parameter estimation methods for Babaurband, Sindh Pakistan*, Energy Conversion and Management, vol. 78, pp.956-967.

Kidmo, D., Danwe, R., Doka, S. and Djongyang, N. (2015) *Statistical analysis of wind speed distribution based on six Weibull Methods for wind power evaluation in Garoua, Cameroon*, Revue des Energies Renouvelables, vol. 18, (2015) 105 – 125, p.21.

Mahmood, F., Resen, A. and Khamees, A., (2020) *Wind characteristic analysis based on Weibull distribution of Al-Salman site, Iraq*, Energy Reports, vol. 6, pp.79-87.

Mohammadi, K., Alavi, O. and McGowan, J. (2017) *Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review*. Energy Conversion and Management, vol. 143, pp.109-122.

Morgan, E., Lackner, M., Vogel, R. and Baise, L. (2011) *Probability distributions for offshore wind speeds*, Energy Conversion and Management, vol. 52(1), pp.15-26.

Moscadelli, M., (2004) *The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee*, SSRN Electronic Journal [Online]. Available at SSRN: https://ssrn.com/abstract=557214 or http://dx.doi.org/10.2139/ssrn.557214 (Accessed10 January 2020).

Okasha, M. and Matter, M., (2015) *On the three parameters Burr type XII distribution and its application to heavy tailed lifetime data*, Journal of advances in mathematics, vol. 10, No.4(ISSN 2347-1921), p.15.

Ouarda, T. B., Charron, C., and Chebana, F. (2016) *Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study*, Energy Convers Manage, vol. 124, pp. 247–65.

Ouarda, T. B., Charron, C., Shin, J.Y. , Marpu, P. R., Al-Mandoos, A. H., Al-Tamimi, M. H. Ghedira, H. and, Al Hosary, T. N. (2015) *Probability distributions of wind speed in the UAE*, Energy Convers Manage, vol. 93, pp. 414–34.

Ozay, C., and Celiktas. M. (2016) *Statistical analysis of wind speed using two-parameter Weibull distribution in Alaçatı region*, Energy Convers Manage, vol. 12, pp. 49–54.

Pfanzagl, J. and Hamböker, R., (1994) *Parametric Statistical Theory*. 1st edition. Edited by Johann Pfanzagl, and Ralf Hamböker. Berlin: De Gruyter.

Pham, H., (2007) *Springer Handbook of Engineering Statistics*. 1st edition. Edited by Hoang Pham. New York: Springer.

Philippe, D. (2015) *Surface Wind-Speed Statistics Modelling Alternatives to the Weibull Distribution and Performance Evaluation*, Boundary-Layer Meteorol, vol. 157, pp. 97–123.

Pobocikova, I. and Sedliackova, Z. (2014) *Comparison of four methods for estimating the Weibull distribution parameters*, Applied Mathematical Sciences, vol. 8, pp.4137-4149.

Sarkar, S., Datta, D., Amit, V., Roy, R. and Phanikanth, V. (2019) *Weibull and Generalized Extreme Value Distributions for Wind Speed Data Analysis of Some Locations in India*, Journal of Civil Engineering, vol. 23(8), PP. 3476-3492.

Shittu, O. and Adepoju, K. (2014) *On the Exponentiated Weibull Distribution for Modeling Wind Speed in South Western Nigeria*, Journal of Modern Applied Statistical Methods, vol. 13(1), pp.431-445.

Shu, Z., Li, Q. and Chan, P. (2015) *Statistical analysis of wind characteristics and wind energy potential in Hong Kong*, Energy Conversion and Management, vol. 101, pp.644-657.

Teimouri, M., Hoseini S. and Nadarajah, S., (2013) *Comparison of estimation methods for theWeibull distribution*, A Journal of Theoretical and Applied Statistics, vol. 47, No. 1, 93–109.

Wright, M., Winter, I., Forster, J. and Bleeck, S., (2014) *Response to best-frequency tone bursts in the ventral cochlear nucleus is governed by ordered inter-spike interval statistics*, Hearing Research, vol. 317, pp.23-32.

Xiao, Y., Li, Q., Li, Z., Chow, Y. and Li, G., (2006) *Probability distributions of extreme wind speed and its occurrence interval*, Engineering Structures, vol. 28(8), pp.1173-1181.

Zaharim, A., Razali, A., Abidin, R. and Sopian, K. (2009) *Fitting of Statistical Distributions to Wind Speed Data in Malaysia*, European Journal of Scientific Research, vol. 26 No.1 (2009), pp.6-12.